

# **High Stakes Lies: Identifying and Using Cues to Deception and Honesty in Appeals for Missing and Murdered relatives**

Thesis submitted in accordance with the requirements of the University of Liverpool  
for the degree of Doctor of Philosophy by Clea Whelan.

May, 2013

# Table of Contents

	<i>Page number</i>
<b>List of Tables</b>	xii
<b>List of Appendices</b>	xv
<b>Abstract</b>	1
<b>Acknowledgements</b>	3

## **Part I: Introduction and Literature Review**

<b>Chapter 1</b>	<b>Deception: Definitions, Meta-analytic Findings, and Dominant Research Methodology</b>	<b>5</b>
<b>1.1.</b>	Definitions of Deception	6
<b>1.2.</b>	Meta-analytic Findings	7
<b>1.2.1.</b>	Cues to Deception	7
<b>1.2.2.</b>	Accuracy in Detecting Deception	9
<b>1.3.</b>	Dominant Methodology in Deception Studies	10
<b>1.3.1.</b>	Limitations of Laboratory-Based Research	11
<b>1.3.2.</b>	Moderating Factors on Cues to Deception	13
<b>1.3.3.</b>	Raising the Stakes in Deception Experiments	14
<b>1.3.4.</b>	Multiple-Cue Approach to Deception Detection	14
<b>1.4.</b>	Conclusions	15

<b>Chapter 2</b>	<b>Cues to Deception in High Stakes Contexts:</b>	
	<b>Theoretical and Methodological Issues</b>	<b>17</b>
<b>2.1.</b>	Major Theoretical Explanations of the Production of Cues to Deception	17
<b>2.1.1.</b>	Cognitive Load	19
<b>2.1.2.</b>	Affective Responses	21
<b>2.1.3.</b>	Behavioural Control and Impression Management	23
<b>2.1.4.</b>	Physiological Arousal	23
<b>2.2.</b>	Research Investigating Cues to Deception in Real Life, High Stakes Situations	25
<b>2.2.1.</b>	Verbal and Paraverbal Cues to Deception	26
<b>2.2.2.</b>	Nonverbal Cues to Deception	29
<b>2.2.3.</b>	Subjective Measures as Cues to Deception	31
<b>2.3.</b>	The Importance of Context	33
<b>2.4.</b>	Conclusions	37
<b>Chapter 3</b>	<b>Accuracy in Detecting Deception in High Stakes Contexts</b>	<b>38</b>
<b>3.1.</b>	The Effect of Ecological Validity on Accuracy	38
<b>3.2.</b>	Accuracy and Confidence in Veracity Judgements of High Stakes Lies	41

<b>3.3</b>	Individual Variation in Accuracy in Detecting Deception	42
<b>3.4.</b>	Consensus Judgements	45
<b>3.5</b>	Conclusions	46

## **Part II: Empirical Research**

<b>Chapter 4</b>	<b>Introduction to the Empirical Research</b>	<b>49</b>
<b>4.1.</b>	Investigating Naturally Occurring, High Stakes Deception	50
<b>4.2.</b>	Investigating the Specific Context of Public Appeals	51
<b>4.3</b>	Research Aims and Questions Relating to the Identification of Cues to Deception	55
<b>4.4.</b>	Research Aims and Questions Relating to Accuracy in Detecting Deception	57
<b>4.5.</b>	Organisation of the Studies	58
<b>4.5.1.</b>	Study 1	59
<b>4.5.2.</b>	Study 2	59
<b>4.5.3.</b>	Study 3	59
<b>4.5.4.</b>	Study 4	60
<b>4.5.5.</b>	Study 5	60
<b>4.5.6.</b>	Study 6	61
<b>4.5.7.</b>	Study 7	61
<b>4.6.</b>	Ethical Considerations	61

<b>Chapter 5</b>	<b>Methodology: The Stimulus Videos</b>	<b>63</b>
<b>5.1.</b>	<b>The Stimulus Materials</b>	<b>63</b>
<b>5.1.1.</b>	<b>Ground Truth</b>	<b>67</b>
<b>5.2.</b>	<b>Between Subjects Stimulus Material Design</b>	<b>71</b>
<b>5.3.</b>	<b>Limitations of the Stimulus Materials</b>	<b>72</b>
<b>Chapter 6</b>	<b>Study 1: A Qualitative Analysis of Cues to Deception Used by Participants Accurate at Detecting Deception in Appeals</b>	<b>74</b>
<b>6.1.</b>	<b>Introduction</b>	<b>74</b>
<b>6.2.</b>	<b>Method</b>	<b>77</b>
<b>6.2.1.</b>	<b>Participants</b>	<b>77</b>
<b>6.2.2.</b>	<b>Materials and Procedure</b>	<b>78</b>
<b>6.3.</b>	<b>Results and Discussion</b>	<b>82</b>
<b>6.3.1.</b>	<b>Deceptive Appealers</b>	<b>83</b>
6.3.1.1.	Fake emotion	86
6.3.1.2.	Lack of emotion	87
6.3.1.3.	Distancing	89
6.3.1.4.	Implausibility	92
6.3.1.5.	Negative personal reaction	93
<b>6.3.2.</b>	<b>Honest Appealers</b>	<b>94</b>
6.3.2.1.	Genuine sadness	94

6.3.2.2.	Containing emotion	96
6.3.2.3.	Genuine/heartfelt	97
6.3.2.4.	Sympathetic personal reaction	98
<b>6.3.3.</b>	<b>Mistakes</b>	<b>99</b>
6.3.3.1.	Emotion not genuine	100
6.3.3.2.	Lack of emotion	100
<b>6.4.</b>	<b>General Discussion</b>	<b>100</b>
<b>Chapter 7</b>	<b>Studies 2 and 3: Quantitative Investigation of Cues to Deception Generated in Study 1</b>	<b>106</b>
<b>7.1.</b>	<b>Introduction</b>	<b>106</b>
<b>7.2.</b>	<b>Study 2: Subjective Observer Ratings of Cues to Deception</b>	<b>108</b>
<b>7.2.1.</b>	<b>Subjective Observer Ratings: Method</b>	<b>109</b>
7.2.1.1.	Participants	109
7.2.1.2.	Materials and procedure	110
<b>7.2.2.</b>	<b>Subjective Observer Ratings: Results and Discussion</b>	<b>111</b>
7.2.2.1.	Inter-rater reliability	111
7.2.2.2.	Analyses of cues rated by observers	111
7.2.2.3.	Case by case analyses using subjective cues rated by observers	117
7.2.2.4.	Accuracy of observers' credibility judgements	118
<b>7.2.3.</b>	<b>Summary of Study 2</b>	<b>120</b>

<b>7.3.</b>	<b>Study 3: Frequency Counts of Verbal Cues to Deception</b>	<b>120</b>
<b>7.3.1.</b>	<b>Verbal Frequency Counts: Method</b>	<b>121</b>
7.3.1.1.	Materials and Procedure	121
<b>7.3.2.</b>	<b>Verbal Frequency Counts: Results and Discussion</b>	<b>123</b>
7.3.2.1.	Analyses of verbal cues using frequency counts	123
7.3.2.2.	Case by case analyses using frequency counts of verbal cues	126
<b>7.3.3.</b>	<b>Combined Case by Case Analyses Using Subjective Observer Ratings and Frequency Counts of Verbal Cues</b>	<b>127</b>
<b>7.3.4.</b>	<b>Verbal Frequency Counts: Summary</b>	<b>130</b>
<b>7.4.</b>	<b>General Discussion</b>	<b>131</b>
<b>Chapter 8</b>	<b>Study 4: Subjective Observer Ratings of Cues to Deception in Appeals</b>	<b>140</b>
<b>8.1.</b>	<b>Introduction</b>	<b>140</b>
<b>8.2.</b>	<b>Method</b>	<b>141</b>
8.2.1.	Participants	141
8.2.2.	Materials and Procedure	141
<b>8.3.</b>	<b>Results</b>	<b>143</b>
8.3.1.	Inter-rater Reliability	143
8.3.2.	Analyses of Cues	144
8.3.3.	Case by Case Analyses	148

<b>8.3.4.</b>	<b>Cues as Measures of Four Factors</b>	<b>152</b>
8.3.4.1.	Factor: fake emotion	152
8.3.4.2.	Factor: norm violation	153
8.3.4.3.	Factor: personal reaction	154
8.3.4.4.	Factor: genuine emotion	155
8.3.4.5.	Case by case analyses using four factors	156
<b>8.3.5.</b>	<b>Accuracy of Observers' Credibility Judgements</b>	<b>159</b>
<b>8.4.</b>	<b>Discussion</b>	<b>160</b>

<b>Chapter 9</b>	<b>Study 5: Frequency Counts of Verbal Cues to</b>	
	<b>Deception in Appeals</b>	<b>164</b>
<b>9.1.</b>	<b>Introduction</b>	<b>164</b>
<b>9.2.</b>	<b>Method</b>	<b>165</b>
9.2.1.	Materials and Procedure	165
<b>9.3.</b>	<b>Results</b>	<b>168</b>
9.3.1.	Analyses of Cues	168
9.3.2.	Case by Case Analyses	173
<b>9.4.</b>	<b>Discussion</b>	<b>177</b>



<b>Chapter 10</b>	<b>Study 6: Frequency Counts of Nonverbal Cues to</b>	
	<b>Deception in Appeals</b>	<b>182</b>
<b>10.1.</b>	Introduction	182
<b>10.2.</b>	Method	184
	<b>10.2.1.</b> Materials and Procedure	184
<b>10.3.</b>	Results	186
	<b>10.3.1.</b> Analyses of Cues	186
	<b>10.3.2.</b> Case by Case Analyses	188
	10.3.2.1. Verbal and nonverbal cues combined	188
	10.3.2.2. Verbal, nonverbal and subjective cues	
	Combined	191
	10.3.2.3. Comparing the classification accuracy	
	of the different groups of cues	196
<b>10.4.</b>	Discussion	197
<b>Chapter 11</b>	<b>Study 7: Police and Public Observer Accuracy in</b>	
	<b>Detecting Deception in Appeals</b>	<b>202</b>
<b>11.1</b>	Introduction	202
	<b>11.1.1.</b> Review of Aims and Hypotheses	205
<b>11.2.</b>	Method	209

<b>11.2.1.</b>	Participants	209
<b>11.2.2.</b>	Materials and Procedure	209
<b>11.3.</b>	Results	213
<b>11.3.1.</b>	Accuracy in Detecting Deception, and Truth Bias	213
<b>11.3.2.</b>	Accuracy and Confidence in Credibility Judgements	216
<b>11.3.3.</b>	Relationship of Consensus Judgements to Veracity	220
<b>11.3.4.</b>	Relationship Between Correct Veracity Judgements and ‘hits’ on Cues	225
<b>11.3.5.</b>	Relationship Between Observer Accuracy and Observer Cue Use.	226
<b>11.4.</b>	Discussion	228

### **Part III: General Discussion and Conclusions**

<b>Chapter 12</b>	<b>General Discussion and Conclusions</b>	238
<b>12.1.</b>	Cues to Deception: Methodological Considerations	239
<b>12.1.1.</b>	Contextual Focus	240
<b>12.1.2.</b>	Between-Subjects Design	242
<b>12.1.3.</b>	Using Cues Generated by Accurate Detectors	242
<b>12.2</b>	Cues to Deception: Findings	243
<b>12.2.1.</b>	Findings in Line With Previous Research	244
<b>12.2.2.</b>	Previously Unidentified Cues	248

<b>12.2.3. Multiple-Cue Approach and Case by Case Analyses</b>	252
<b>12.3. Theoretical Development</b>	253
<b>12.3.1. The Social Interactionist (SI) Account</b>	255
<b>12.3.2. The Individual Behavioural Profile (IBP) Account</b>	259
<b>12.4. Observer Accuracy in Detecting Deception</b>	261
<b>12.4.1. General Accuracy</b>	262
<b>12.4.2. Police vs Non-Police: Accuracy, and Truth and                     Deception Bias</b>	264
<b>12.4.3. Relationships Between Accuracy and Confidence</b>	266
<b>12.4.4. Consensus Judgements as a Predictor of Veracity</b>	267
<b>12.4.5. Relationships Between Consensus and Cues</b>	268
<b>12.5. Methodological Limitations</b>	269
<b>12.6. Suggestions for Future Research</b>	274
<b>12.7. Practical Applications and Implications</b>	277
<b>12.8. Conclusions</b>	280
 <b>References</b>	 282
 <b>Appendices</b>	 301

## List of Tables

<b>Table 5.1</b>	Properties of appeals used as stimulus materials in studies 1, 2 and 3	64
<b>Table 5.2</b>	Properties of appeals used as stimulus materials in studies 4, 5 and 6	66
<b>Table 5.3</b>	Properties of appeals used as stimulus materials in Study 7	67
<b>Table 5.4</b>	Frequency of case evidence used to establish ground truth	69
<b>Table 6.1</b>	Qualitative analysis; themes and their composite categories (number of units relating to category in brackets), with examples	84
<b>Table 7.1</b>	Kendall's <i>W</i> s, means, <i>SD</i> s, <i>CI</i> s, <i>F</i> ratios and <i>p</i> values, and effect sizes for all subjective cues rated by observers	114
<b>Table 7.2</b>	Inter-rater reliability, means, ( <i>SD</i> s), <i>CI</i> s, <i>F</i> ratios and <i>p</i> values, and effect sizes for frequency counts of each verbal cue	125
<b>Table 7.3</b>	Case by case analyses: cues which discriminated significantly between honest and deceptive appeals: overall cue mean, score per case (appear), and overall 'hit' rate. For each case: 'hit' rate of subjective cues, 'hit' rate of verbal frequency count cues, and 'hit' rate of combined cues. All 'hit's related to overall cue means.	129

<b>Table 8.1</b>	Kendall's <i>W</i> s for cues rated subjectively by observers; means, <i>SD</i> s, <i>CI</i> s, <i>F</i> ratios and <i>p</i> values, and effect sizes for cues which achieved significant inter-rater reliability	146
<b>Table 8.2</b>	Case by case analyses; cues rated by observers which discriminated significantly between truthful and deceptive appeals. Overall cue mean, score per case (appealer), and overall 'hit' rate. For each case: 'hit' rate of cues. All 'hit's related to overall cue means.	149
<b>Table 8.3</b>	Correlations (with significance values) for cues in 'fake emotion' factor ( <i>N</i> = 32)	153
<b>Table 8.4</b>	Correlations (with significance values) for cues in 'norm violation' factor ( <i>N</i> = 32)	154
<b>Table 8.5</b>	Correlations (with significance values) for cues in 'personal reaction' factor ( <i>N</i> = 32)	155
<b>Table 8.6</b>	Correlations (with significance values) for cues in 'genuine emotion' factor ( <i>N</i> = 32)	156
<b>Table 8.7</b>	Case by case analyses: factor mean, score per case (appeal), and overall 'hit' rate. For each case: overall 'hit' rate of factors. All 'hit's related to overall cue means	158
<b>Table 9.1</b>	Inter-rater reliability, percentage means ( <i>SD</i> s), <i>CI</i> s, <i>F</i> ratios and <i>p</i> values, and effect sizes for verbal cues	172
<b>Table 9.2</b>	Case by case analyses using verbal cues: score per case (appeal), and overall 'hit' rate. For each case: overall 'hit' rate of cues. All 'hit's related to overall cue means	175

<b>Table 10.1</b>	Inter-rater reliabilities, means ( <i>SDs</i> ), <i>CI</i> s, <i>F</i> ratios and <i>p</i> values, and effect sizes for frequency counts of nonverbal cues as a proportion of the total appeals	188
<b>Table 10.2</b>	Case by case analyses using verbal and nonverbal cues: score per case (appeal), and overall ‘hit’ rate. For each case: overall ‘hit’ rate of cues. All ‘hit’s related to overall cue means	190
<b>Table 10.3</b>	Case by case analyses; verbal cues, nonverbal cues, and cues rated by observers. Overall cue mean, score per case (appeal), and overall ‘hit’ rate. For each case: ‘hit’ rate of cues. All ‘hit’s related to overall cue means.	193
<b>Table 10.4</b>	Classifications (correctly classified, misclassified, and unclassified) of appeals using different groups of cues. Classifications shown as percentages, followed by number of cases classified in brackets	196
<b>Table 11.1</b>	Accuracy of observers’ credibility judgements of appeals	216
<b>Table 11.2</b>	Observers’ self-estimates of pre & post test accuracy, and confidence rated on a 5 point scale, in their credibility judgements of appeals	217
<b>Table 11.3</b>	For each appeal; percentage of accurate judgements (overall, by police, and by the public), and percentage of ‘hits’ on cues identified in Studies 4, 5 and 6 as discriminating between honest and deceptive appeals	222

## List of Appendices

<b>Appendix 1</b>	Participant information sheets	301
<b>Appendix 2</b>	Participant debrief sheets	305
<b>Appendix 3</b>	Case summaries of all appeals used as stimulus materials	307
<b>Appendix 4</b>	Study 2: participant response sheet	334
<b>Appendix 5</b>	Study 3: transcripts of appeals	341
<b>Appendix 6</b>	Study 3: definitions and instructions to coders for frequency counts of verbal cues in appeal transcripts	346
<b>Appendix 7</b>	Study 4: participant response sheet	348
<b>Appendix 8</b>	Study 5: transcripts of appeals	352
<b>Appendix 9</b>	Study 5: definitions and instructions to coders for frequency counts of verbal cues in appeal transcripts	369
<b>Appendix 10</b>	Study 6: Definitions of nonverbal cues and instructions to coders	372
<b>Appendix 11</b>	Study 7: participant instructions and response sheet	373

## **Abstract**

Deception is a near-universal human behaviour, and decades of psychological research have generated hundreds of studies investigating behaviours related to deception (often referred to as ‘cues to deception’), and accuracy in detecting deception. However, the low ecological validity of most of this research on deception is a limitation that has been recognised by many authorities in the field, who have argued that, to produce findings that are relevant to real world, high stakes deceptive behaviour, there is a need for more research using real life, high stakes lies. Consequently, the research presented investigated cues to deception using video footage of people making real life, public appeals for help with missing or murdered relatives. There were four major aims of the research; to develop and apply a methodology to identify cues to deception in a specific high stakes context, and to investigate whether these cues were useful in discriminating between honest and deceptive appeals in this context; to develop a multiple-cue approach to behaviours which discriminate between honesty and deception in the context of appeals, and which may be of applied use; to develop a theoretical rationale for cues that may be context-specific and may emerge only in high stakes contexts, and cannot readily be accommodated by existing theoretical approaches; and to investigate the effect of using ecologically valid stimulus materials on the accuracy of observers (both police and non-police) in detecting deception. To these ends, seven studies were conducted. Study 1 was a qualitative investigation of cues to deception in appeals used by people unusually accurate at detecting deception in the context of appeals. The aim of the study was to generate cues, including possibly new and context-specific cues, to be tested in future, quantitative research. Studies 2, 3, 4, 5 and 6 were then



quantitative investigations of nonverbal cues, verbal cues, and subjective observer ratings identified in Study 1 and in previous research on high stakes deceptive behaviour, in the context of appeals. Finally Study 7 was an investigation of observer accuracy in detecting deception, truth bias, accuracy-confidence relationships, and the value of consensus judgements in the high stakes context of appeals, including samples of police and non-police participants. Several important findings emerged. In the final analyses, 23 cues discriminated between honest and deceptive appeals, the majority of which had not previously been identified, and may, at least to some degree, be context specific. These included two nonverbal cues, four verbal cues, and 17 subjective measures. Utilising a multiple-cue approach, case by case analyses combining information from the verbal and nonverbal cues correctly classified 78% of individual cases. Observer accuracy in discriminating between honest and deceptive appeals was above chance ( $M = 71\%$ ), and police observers ( $M = 72\%$ ) were more accurate than non-police observers ( $M = 68\%$ ). Though non-police observers demonstrated a truth bias, police observers showed some evidence of a deception bias. There was also a positive relationship between confidence and accuracy for all groups of observers. Importantly, observer consensus using simple global subjective judgments of veracity correctly predicted 92% of cases. To account for the findings regarding the previously unidentified cues, various theoretical approaches are considered, including two new developments, the social interactionist approach, and the individual behavioural profile approach. Various implications for research and practice are discussed, including the possible development of consensus judgments, and a checklist of cues, for use in police investigations in this type of context.

## **Acknowledgements**

My most heartfelt gratitude goes to my supervisor, Prof. Graham Wagstaff, for his unflagging patience, wisdom, generosity, encouragement, and above all his good humour throughout this project. My thanks also go to Dr Jacqueline Wheatcroft, for her advice and support. An important part of this thesis would not have been possible without the participation of Merseyside police officers, and I would like to thank those who participated for giving up their time, and also Prof. Jon Cole, for introducing me to the police officers who enabled this.

My family and friends know how much their love, belief, support, and encouragement has meant to me.

# **Part I**

## **Introduction and Literature Review**

# **Chapter 1**

## **Deception: Definitions, Meta-analytic Findings, and Dominant Research Methodology**

Deception is a near-universal human behaviour; research indicates that, on average, people lie between one and two times a day (DePaulo, Kashy, Kirkendol, Wyer & Epstein, 1996; Hancock, Thom-Santelli & Ritchie, 2004; Serota, Levine & Boster, 2010). Some researchers have even suggested that deception is not restricted to our own species, but that, for example, orang-utangs, chimpanzees, and capuchin monkeys, may also exhibit deceptive behaviour (Gretschel, Haun, Liebel & Kaminski, 2012; Hare, Call & Tomasello, 2006; Wheeler, 2009). Given this ubiquity, it is not surprising that deception has been a subject of investigation in disciplines as various as psychology, philosophy, linguistics and psychiatry. In particular, decades of psychological research have generated hundreds of studies investigating behaviours related to deception (often referred to as ‘cues to deception’), and accuracy in detecting deception (both human and automated).

## **1.1 Definitions of Deception**

The Collins English Dictionary (2010) defines ‘deceive’ as ‘to mislead by deliberate misrepresentation or lies’, and definitions of deception used in psychological research have expanded on this basic theme. Researchers (for example, Buller & Burgoon, 1996; Ekman, 2001; Vrij, 2008) generally agree that three criteria must be met to classify communication as deceptive. First, deceivers must believe that their communication is not true. Second, deception must be an intentional act, so that, for example, forgetting, misremembering, mistaking, self-deceiving, or misperceiving are not deceptive. Third, the receiver of the deceptive communication must have no prior warning that he or she is being deceived.

Deception can be conceptualised as a control of information in order to manipulate the beliefs, or psychological or cognitive states of others (Buller & Burgoon, 1996; McCornack, 1992), and it has been suggested that the purpose of deception is to hamper the decision-making abilities of the receiver of the deceptive communication (Wardle & Gloss, 1982). Ekman (2001), and Turner, Egdley and Olmstead (1975) argue that this manipulation of information occurs in two main forms, concealment, and falsification (although Turner et al. also suggest that diversionary responses may function deceptively). Concealment involves withholding true information, and falsification involves the presentation of false information as if it were true. A deceptive communication may involve both concealment and falsification, and as DePaulo et al. (2003) have noted, deceptive accounts may also contain some truth, leading some researchers to suggest that deception is not opposite to truth, but rather exists on a continuum (Rycyna, Champion & Kelly, 2009). However, as it is the intention of the deceiver, rather than

the specific form of the deception, that is generally agreed on by researchers as being a necessary criterion for the classification of a communication as deceptive, deception is almost always investigated as a dichotomous construct. In light of the above, Vrij's (2008) definition is perhaps most suited to the purposes of the present thesis: i.e. 'a successful or unsuccessful attempt, without forewarning, to create in another a belief which the communicator considers to be untrue' (p.15).

## **1.2 Meta-analytic Findings**

Meta-analyses of several decades of research into deception have produced two major findings; these are that cues to deception tend to be unreliable and/or weak (DePaulo et al., 2003), and, in general, people are poor at detecting deception (Bond & DePaulo, 2006).

### **1.2.1. *Cues to Deception***

In DePaulo et al.'s (2003) meta-analysis of cues to deception, no 'Pinocchio's nose' emerged that unfailingly discriminates between liars and truth-tellers; of the 158 behaviours examined from 116 reports, the duration and/or frequency of only 24 behaviours were found to be consistently related to deception across studies. According to the authors' organisation of these cues, the largest group of cues was associated with the idea that liars tell less compelling tales than truth tellers. Thus liars were found to produce more verbal and vocal uncertainty, chin raises, and word

and phrase repetitions, and to be more discrepant or ambivalent than truth tellers. Liars were also found to show or produce less verbal and vocal immediacy, less involvement, fewer illustrators, and their tales were perceived as less plausible and less logical than those of truth tellers. Some cues suggesting that liars were more tense than truth tellers also emerged as significant; thus liars were found to produce more pupil dilation, fidgeting, and vocal tension, to have higher vocal pitch, and to be more nervous and tense overall, than truth tellers. Interestingly, three cues also emerged that suggested that liars were perceived to be less pleasant than truth tellers; thus liars were found to produce more negative statements and complaints, and to be less cooperative and to appear less facially pleasant than truth tellers. Finally, three cues indicated that lies included fewer ordinary imperfections and unusual contents than truths; thus lies were found to contain fewer spontaneous corrections and admitted lack of memory, and more related external associations, than truths. Importantly, De Paulo et al. (2003) note that honesty and deception may be discriminated more powerfully using subjective measures, rather than objective measures.

In general, however, although some cues emerged as discriminatory in the meta-analysis, the large majority did not, suggesting that findings regarding cues to deception tend to be difficult to replicate consistently. Indeed, of the 24 cues discussed above, 10 were based on only a small number of estimates (five or less). Furthermore, effect sizes were generally small, with a median of  $d = 0.10$ ; even the most diagnostic cue (cooperativeness) had only a moderate effect size of  $d = 0.66$ . As the differences between liars and truth tellers are usually small and thus are likely to be difficult to discern, these faint cues may be difficult to use in real life contexts. However, the authors suggested several ways in which the results of their analyses

may have underestimated the potential for cues to separate truths from lies, three of which are relevant to the present thesis. First, as in most studies in the literature, the predictive power of each cue was tested individually; however, it could be the case that a consideration of combinations of cues may increase the degree to which lies can be discriminated from truths. Second, some cues may be relevant only within a particular context, and may not be generalisable to other situations. Third, as previously noted, truths and lies may be discriminated more powerfully by using subjective measures rather than objective ones.

### **1.2.2 *Accuracy in Detecting Deception***

As noted, meta-analytic findings suggest that, in general, people are poor at detecting deception. Bond and DePaulo (2006), in their meta-analysis, report an overall average accuracy in detecting deception of 54%, which is barely above the 50% accuracy that would be expected by chance. The authors also report that people are more able to correctly identify truths (61% accuracy) than lies (47% accuracy), indicating a truth bias; that is, observers are more likely to judge a communication as honest than deceptive, and thus truthful messages are correctly identified more frequently than deceptive messages. Explanations usually offered for this effect derive from the availability heuristic; i.e. in daily life, people generally encounter more truthful than deceptive messages, resulting in an a priori expectation of truthfulness and a cognitive bias towards believing communications (Burgoon, Blair & Strom, 2008; O'Sullivan, Ekman & Friesen, 1988). Furthermore, a review of the research has suggested that there is no relationship between accuracy and confidence



in making judgements of veracity (DePaulo et al., 1997), such that when people attempt to detect deception, they report feeling as confident in their decisions when they are incorrect, as when they are correct.

Nevertheless, some researchers have reported that some groups and individuals may be better than others at detecting deception. For example, Ekman, O'Sullivan and Frank (1999) found that federal officers (mostly from the Central Intelligence Agency), county sheriffs, and clinical psychologists interested in deception were better able than others to detect lies (with accuracy rates of 73%, 67%, and 68% respectively). O'Sullivan and Ekman (2004), have tested thousands of individuals from a wide variety of professions and identified a small number who were unusually accurate at detecting deception, and Bond (2008) reports identifying two individuals, both correctional officers, who scored above 80% accuracy on both parts of a repeated deception detection task; though, like others, they report that accurate deception detection does not appear to be the norm. It would obviously make sense to see these meta-analytic findings as related; that is, if the differences in behaviours between liars and truth-tellers are often inconsistent and small, they are likely to be difficult to detect in real life, and people will generally be unable to detect deceit. However, it is also possible that such findings may reflect some major limitations in the dominant methodology used in deception research.

### **1.3 Dominant Methodology in Deception Studies**

In DePaulo et al.'s (2003) meta-analysis it was reported that out of 116 studies, 101 used university students as participants, and most were conducted in laboratory

settings. The poor ecological validity of the large majority of deception studies is a limitation that has been recognised by a number of leading researchers in the field (for example, DePaulo & Morris, 2004; Frank & Svetieva, 2012; Granhag & Stromwall, 2004; Porter & ten Brinke, 2008; Shuy, 1998; Vrij, 2004). These researchers have argued that there are very likely to be differences between cues exhibited in low stakes situations, and those exhibited in high stakes situations. Therefore, although the findings of most previous studies may be applicable to the low stakes situations in which they were carried out, they may have less relevance to more high stakes situations, such as most forensic contexts. The problems may be identified as follows.

### ***1.3.1. Limitations of Laboratory-Based Research***

Laboratory experiments are necessarily limited by ethical considerations, and participants are, therefore, generally required to tell low stakes, ‘white’ lies; these lies cannot be of a serious nature, with serious consequences, and so participants are generally instructed to lie about something trivial. Furthermore, there is little motivation for participants to succeed in the lie (though in some studies, a low financial reward may be offered), as the participants know they are taking part in an experiment, and participants are usually instructed to lie by the researcher and so are not responsible for their lies. Examples of lies that participants have been required to tell in this type of study include, lying about who they saw erasing information from a blackboard, lying about a recently viewed film, lying about autobiographical events, and lying about details in a cartoon story (Cohen, Beattie & Shovelton, 2010;

Porter et al., 2008; Vrij, Akehurst, Soukara & Bull, 2004; Vrij, Edward, Roberts & Bull, 2000). These types of low stakes lies may differ from more high stakes lies; as when, for example, lying about committing a crime, in which the consequences of not being believed are serious for the person telling the lie.

One implication is that the use of low stakes situations may lead to underestimates of the effectiveness of some cues in detecting deception in more realistic forensic situations. For example, in line with above, Miller and Stiff (1993) have suggested that in typical laboratory studies, the stakes are not high enough to elicit discernible cues to deception, and thus deceit is almost impossible to detect. In forensic contexts in particular, lies are likely to be high stake, as the consequences for the liar of not being believed may be severe, and these high stakes may be difficult, if not impossible, to replicate in the lab. Furthermore, experimental, laboratory-based studies, in which participants are randomly allocated to groups, do not allow for possible individual differences to emerge in the type of person who may choose to deceive, particularly in high stakes situations.

Nevertheless, despite the limitations of laboratory experiments, there has been very little research investigating cues to deception in real life, high stakes situations (see, for example, Davis, Markus, Walters, Vorus & Connors, 2005; Haprstter, Adams & Jarvis, 2009; ten Brinke & Porter, 2012; Vrij & Mann, 2001a), and it is to a large extent unknown how much relevance the cues to deception identified in low stakes studies, have to high stakes contexts. This is an important limitation, as it is in high stakes contexts, particularly in forensic contexts, in which the consequences of faulty credibility judgements may be most serious. The few studies in the literature which have attempted to address this issue will be discussed in Chapter 2.

### ***1.3.2. Moderating Factors on Cues to Deception***

Further relevant here is the proposal by DePaulo and Morris (2004) that highly charged situations in which motivation is strong, and consequences are serious, may produce more reliable cues to deception. In their meta-analysis, DePaulo et al. (2003) found that strong motivation to succeed in the lie, lies which were identity relevant, and lies about transgressions, were all moderating factors that produced more prominent cues to deception, even in relatively low stake situations. These are factors that are more likely to be present in high stakes situations than in low stakes situations, particularly in a forensic context. DePaulo et al. (2003) define a transgression as a behaviour regarded as serious and which elicits punishment or disapproval, and identity relevant lies as those which protect the liars' image to conceal a discreditable deed; one might expect both these factors to be important in a forensic context, along with strong motivation to succeed in the lie in order to escape punitive consequences. Correspondingly, one might expect high stakes situations to produce more prominent and reliable cues to deception. Supporting this suggestion, Bond and DePaulo (2006) report that accuracy in discriminating lies from truths is increased when judging motivated, rather than unmotivated 'senders'. It seems likely, therefore, that the use of real life, high stakes lies in deception studies will facilitate the emergence of cues to deception, and the accuracy of observers in detecting deception.

### ***1.3.3. Raising the Stakes in Deception Experiments***

With these issues in mind, some efforts have been made by researchers to develop paradigms with greater ecological validity by using antisocial deception. For example, experiments have been designed to elicit unsanctioned lying about cheating on an experimental task (Blair, Levine & Shaw, 2010; Feeley & deTurck, 1998), and participating in mock crimes (Hartwig, Granhag, Stromwall & Kronkvist, 2006; Kassin & Fong, 1999). However, it is not clear how closely these paradigms actually mirror the motivation involved in, and the consequences of, deception in forensic contexts. In the unsanctioned cheating paradigm, students are faced with the prospect of academic staff being informed of their cheating on an experimental task, which doubtless raises the stakes above the usual level of deception experiments. However, it is doubtful whether social disapproval is a consequence equivalent to the possible punitive consequences of lying in a high stakes forensic context. In the mock crime paradigm, participants are instructed to commit the crime and to lie by the experimenter, and so even in studies in which participants are interviewed by police, they are aware that the deception will have no serious consequences.

### ***1.3.4. Multiple-Cue Approach to Deception Detection***

Another critical issue in deception methodology concerns how cues are used. As previously noted, DePaulo et al. (2003) suggest that a consideration of combinations of cues, rather than individual cues, may increase the degree to which lies can be discriminated from truths. Similarly, several researchers have suggested that the

more alleged deceptive cues that a person exhibits, the more likely that person is to be lying, and hence have advocated a multiple-cue approach to deception detection (for example, ten Brinke & Porter, 2012; Vrij & Mann, 2004). In other words, if there are, indeed, no universal cues to deception, the use of a range of cues to assess individual cases may overcome the problems associated with focusing on single cues. In line with this, studies using this multiple-cue approach (Ekman, O'Sullivan, Friesen & Scherer, 1991; Leal, Vrij, Mann & Fisher, 2010) have found high accuracy rates (86% and 79% respectively) using a case by case method of discrimination by subject rather than looking at differences between groups; i.e. by classifying individual cases as honest or deceptive based on the frequency of several cues, rather than looking at the frequencies of occurrence of single cues in honest and deceptive groups. However, despite these promising results, this is an approach that has rarely been adopted in deception research.

#### **1.4 Conclusions**

To summarize, deception research has shown very few useful indicators of deception, and people are generally very poor at detecting whether others are lying or telling the truth. However, it could be the case that these disappointing results may at least partly be a feature of the dominant methodology that has been adopted so far; i.e. the use of low stakes situations of poor ecological validity. It is possible, therefore, that raising the stakes of the deception involved may increase the production and salience of cues to deception, and, as a corollary, this is likely to have an effect on the ability of individuals to detect deception, as successful detection of

deception is dependent on cues. Also, the use of a multiple-cue approach may also increase the ability to detect deception, as it acknowledges the fact that single cues tend not to be robust across situations. In view of this, the present thesis investigated cues to deception, and accuracy in detecting deception, in a high stakes context, using a multiple-cue approach.

# **Chapter 2**

## **Cues to Deception in High Stakes Contexts: Theoretical and Methodological Issues**

In Chapter One, the key issue of ecological validity in deception studies was introduced. To understand in more detail how raising the stakes may affect the production of cues to deception, it is useful to discuss the major theoretical factors believed to be involved in the production of deceptive behaviours, and how these factors may be affected in high stakes contexts. It was also noted in Chapter One that there are only a handful of studies which have investigated cues to deception in real life, high stakes situations, and this small body of research is discussed in detail below.

### **2.1 Major Theoretical Explanations of the Production of Cues to Deception**

There have been a number of theoretical proposals, with varying degrees of support, which have attempted to address the question of why deceptive behaviour may differ from honest behaviour. For example, interpersonal deception theory (Buller & Burgoon, 1996) takes account of interpersonal communicative processes to predict



behaviours related to deception in interactive contexts. According to this approach, with increased interaction, the liar acquires feedback from the receiver of the lie, enabling him or her to exert greater control over his or her performance. Over the course of the interaction, therefore, liars should display greater composure, greater involvement, and more immediacy in their communication. However, cues related to involvement, immediacy, composure and fluency have generally been found not to vary with interactivity as predicted by the model (DePaulo et al, 2003).

An alternative model, the self-presentational perspective (DePaulo et al., 2003) proposes that, most commonly, lies are low stakes, and people lie primarily to manage impressions and protect identities, rather than for material gain or to escape from punishment. According to this perspective, deceptive self-presentations will not be portrayed and embraced as convincingly as truthful self-presentations; hence, for example, liars will appear less forthcoming, pleasant, and compelling, and more tense than truth-tellers. Also, people are expected to exhibit more evidence of deliberation, and less spontaneity during deceptive communications than during honest communications; so, for example, liars may include fewer ordinary imperfections and unusual details in their communications, than truth-tellers. Again, support for these predictions has been mixed; for example, in relation to liars being less forthcoming than truth-tellers, meta-analytic findings indicate that there is no difference in overall length between honest and deceptive communications, although liars provide fewer details than truth-tellers (DePaulo et al, 2003). Similarly, in relation to the proposal that liars display more deliberation, findings suggest that, although liars have been found to include fewer spontaneous corrections in their accounts than truth-tellers (DePaulo et al., 2003), they also produce more speech errors (Sporer & Schwandt, 2006).

Nevertheless, the self-presentational perspective is now commonly included within the four major factors that are popularly considered by researchers to lie behind the production of cues to deception. These factors are summarized in Zuckerman, DePaulo and Rosenthal's (1981) Four Factor Model as: cognitive load, affective responses, behavioural control and impression management, and arousal.

### ***2.1.1. Cognitive Load***

One of the most widely accepted theoretical propositions in the deception literature is that lying is intrinsically more cognitively demanding than telling the truth (see, for example, DePaulo et al., 2003; Frank & Feeley, 2003; Vrij et al., 2008; Zuckerman et al., 1981). Liars face the complex, simultaneous tasks of suppressing the truth, creating novel material which is consistent with what is already known by the receiver of the lie, and remembering the details of what has already been said in the deceptive communication. Furthermore, as liars typically take their credibility for granted less than truth-tellers (see, for example, DePaulo et al., 2003; Kassin & Gudjonsson, 2004), they may monitor their self-presentation to suppress certain behaviours, role play and simulate certain behaviours in order to appear credible, and monitor the receiver's reactions for signs of incredulity. A number of laboratory experiments have shown that truth-tellers, on the other hand, tend to believe in the visibility of innocence and therefore 'tell the truth like it happened', thus avoiding the cognitive demands associated with deception (Hartwig, Granhag, & Stromwall,

2007; Hartwig, Granhag, Stromwall & Vrij, 2005; Hartwig et al., 2006; Stromwall, Hartwig & Granhag, 2006; Stromwall et al., 2006).

The increase in cognitive load which is expected during acts of deception may result in behaviours related to cognitive load, for example, an increase in speech disruptions and errors, an increase in gaze aversion, and a reduction in illustrators (see, for example, Frank & Feeley, 2003; Zuckerman et al., 1981). Based on this premise, a body of research has recently emerged in which strategies have been developed to increase cognitive load in order to maximise cues to deception related to increased cognitive load. Examples include: recalling an event in reverse order, which results in liars producing slower speech, and more speech hesitations/errors (Vrij et al., 2008); and an instruction to answer questions as quickly as possible, which results in truth-tellers producing quicker response times, fewer inconsistencies and more eye movements than liars, all suggesting lower cognitive loads (Walczyk et al., 2012).

Cues to deception related to cognitive load could be expected to emerge in both high and low stakes situations, as both situations may require cognitive effort. However, the effects of cognitive load may be magnified according to the degree in which the cognitive tasks enumerated above are in effect. For example, the cognitive demands of deception are likely to be higher in situations in which the liar is motivated to be believed, as it is only in these circumstances in which liars may take their credibility less for granted, and so monitor their behaviour and the reactions of the receiver (Vrij, Mann, Leal & Fisher, 2010). Furthermore, cognitive demands are increased in situations in which the liars' knowledge of the truth can be easily and clearly accessed; for example, because it is distinctive or meaningful, so that it is then difficult to suppress the truth (Vrij et al., 2010). Strong motivation to succeed in

the lie, and lying about distinctive or meaningful events, are both factors which are more likely to emerge in high stakes situations, as when, for example, lying about a serious crime, than in low stakes situations.

### **2.1.2. *Affective Responses***

In addition, a number of affective responses have been hypothesised to relate to deception, ranging from guilt, fear, shame, and anxiety, to excitement in duping the receiver of the lie (Ekman, 2001; Zuckerman et al., 1981). It has been argued that these emotions may produce behaviours related to the affective experience, which liars may attempt to conceal, but of which they may nevertheless ‘leak’ evidence (Ekman, 2001). For example, liars experiencing guilt and anxiety may attempt to distance themselves from their own communication, resulting in behaviours related to indirect communication, such as gaze aversion (note that gaze aversion is also associated with cognitive load), and evasive language (Zuckerman et al., 1981). In support of this, meta-analytic findings indicate that liars display more verbal and vocal uncertainty, and less verbal and vocal immediacy, and involvement, than truth-tellers (DePaulo et al., 2003).

Liars may also betray their true emotions with facial expressions. Darwin’s inhibition hypothesis (1872) proposes that some aspects of genuine facial expression cannot be voluntarily simulated, but may be involuntarily expressed when a genuine emotion is felt. Building on this, Ekman (2001) has proposed that, during deception, emotional leakage in facial expressions is an involuntary and near-ubiquitous human behaviour, and, in support, Porter and ten Brinke (2008) found that all of their

participants leaked inconsistent emotional expressions when trying to produce deceptive expressions of emotion.

However, clearly, affective responses may differ between low stakes and high stakes situations, and as deception in high stakes contexts is more likely to be accompanied by strong emotions, the presence of high stakes contexts may be central to the elicitation of cues to deception related to affective responses (Ekman, 2001; Frank & Feeley, 2003; Shaw, Porter & ten Brinke, 2013). For example, high stakes deception such as lying about a serious crime during a police interview, may produce feelings of fear and anxiety, as the consequences of not being believed may be serious. In addition, if the lie is personally relevant to the liar, the liar is more likely to experience guilt or shame (Lazarus, 1991). In comparison, low stakes deception, such as lying about enjoying a film, is unlikely to be accompanied by anxiety and fear of getting caught, or by guilt and shame. Thus, low stakes lies and high stakes lies can be expected to differ in the affective signals generated by the liar (Frank & Feeley, 2003).

Furthermore, the anxiety that may be associated with deception in high stakes situations, may also increase cognitive load. The majority of the tasks identified as increasing cognitive load in liars could be described as ‘frontal’ tasks; i.e. they are likely to involve the frontal areas of the brain typically associated with making planned or ‘executive’ decisions. As such, performance of such tasks is likely to be made more difficult by increases in anxiety and arousal, which may produce symptoms that are intrusive and also ‘overload’ the brain’s executive capacity (Wagstaff, et al., 2008).

### ***2.1.3. Behavioural Control and Impression Management***

As previously noted, the third main factor allegedly underlying cues to deception is the idea that liars may attempt to control their behaviour in an effort to suppress signs of deception, and simultaneously behave in a credible manner (DePaulo et al., 2003; Zuckerman et al., 1981). For example, in addition to showing more deliberation, and less spontaneity and detail, liars may attempt to suppress signs of nervousness, such as fidgeting, and this may paradoxically result in behaviours that appear rigid or inhibited (Vrij, 2004). Attempts to control behaviour may also result in liars being perceived by observers as less cooperative, less pleasant, and less plausible (DePaulo et al., 2003).

As discussed above, high stakes situations may result in strong affective responses which a liar may need to inhibit. According to the impression management approach, in high stakes situations, liars may not only be required to suppress genuinely felt, powerful emotions, they may also be required to simultaneously fabricate displays of emotion to appear credible. For example, a parent publicly pleading for the return of a missing child may be expected to display distress and sadness. This may result in, for example, facial expressions which do not match the emotion that the liar is claiming to experience (ten Brinke & Porter, 2012).

### ***2.1.4. Physiological Arousal***

The fourth and final factor alleged to underlie cues to deception concerns psychological discomfort or excitement produced by acts of deception. The proposal

is that lying may result in increased arousal in the autonomic nervous system, resulting in largely involuntary physiological reactions related to activation of the sympathetic nervous system (Zuckerman et al., 1981). These may include increased perspiration, blood pressure, heart rate, and respiration rates (Koper & Sahlman, 1991); i.e. reactions on which the polygraph is based. There is some support for this proposal. For example, meta-analytic findings (DePaulo et al, 2003) indicate that deception is associated with a higher vocal pitch, which is in line with psychophysiological research showing that arousal is associated with involuntary contraction of the muscles of the vocal tract, resulting in higher pitch (for example, Johnstone & Scherer, 2000). Deception is also associated with pupil dilation (DePaulo et al, 2003), which is a further indication of arousal (for example, Partala & Surakka, 2003), although it should be noted that pupil dilation may also be related to cognitive load (for example, Beatty, 1982).

Importantly, in terms of the present discussion, as physiological arousal is a condition of alertness and readiness in response to stimulation, it might be expected to increase in high stakes situations where the perceived threat of the stimulation is increased; also the discomfort or excitement resulting in arousal might be expected to increase in relation to the consequences of the deception.

It can be noted here that researchers tend to agree that Zuckerman et al.'s (1981) four factors are not mutually exclusive; they may both co-occur and interact, and behaviours may be predicted by more than one factor (for example, DePaulo et al., 2003; Vrij, 2004). Nevertheless, it could be argued that the salience of each factor can be expected to increase as stakes are raised, potentially resulting in more prominent cues to deception. Indeed, importantly, one of the central tenets of the present thesis is that each of these factors may become more pronounced or

influential in high stakes contexts. Lying in high stakes situations may be, cognitively and emotionally, a complex and difficult experience; the liar may be required to simultaneously inhibit and simulate emotions and behaviours, monitor themselves and others in their environment, and fabricate a consistent and convincing account, and to do this with the awareness of serious negative consequences of failure. It would not be surprising if such factors had a number of psychological and physiological consequences.

## **2.2. Research Investigating Cues to Deception in Real Life, High Stakes Situations**

As mentioned previously, the few studies that have investigated cues to deception in real life, high stakes lies, have produced some promising, but mixed, results. These results are discussed below; for clarity, the cues are organised into groups of verbal and paraverbal behaviours, nonverbal behaviours, and subjective measures. Within the literature, verbal cues are normally considered those that relate directly to language used, and, therefore, consist of the words chosen by the sender. Paraverbal cues relate, not to the choice of words, but to how those words are said, and qualities of the voice; they include, for example, elements such as intonation, pitch, volume, speech rate, emphasis and dysfluency. Nonverbal cues are not related to word choice, or how words are said, but to other physical behaviours; so they may include, for example, body language and facial expressions. Subjective measures are qualities of the sender that are measured as subjective impressions of observers.



### ***2.2.1. Verbal and Paraverbal Cues to Deception***

The first major attempt to develop a systematic method of investigating and assessing high stakes deception through an analysis of verbal cues was Statement Validity Analysis (Stellar & Köhnken, 1989). In 1955 the German Supreme Court passed a ruling that in all cases in which child sexual abuse was an issue, psychologists should be called to give evidence as to the reliability of the testimony. By 1982, expert testimony had been given in over 40,000 cases, and various criteria had been drawn up to assess the credibility of witness statements. Based on this work Stellar and Köhnken (1989) compiled a list of such criteria, and developed a procedure to assess the truthfulness of witness statements, known as Statement Validity Analysis (SVA). A major feature of this approach is the use of Criteria Based Content Analysis (CBCA) which is a method of looking at the verbal contents of a statement. For example, according to CBCA, truthful statements are alleged to contain more unstructured production (information scattered but logical), details (place, time, people, objects, events), contextual embedding (information placed in everyday context), descriptions of interactions (three or more parts, e.g. A leads to B doing X leads to A doing X), reproduction of speech or conversation, unexpected complications (e.g. event interrupted by someone else, phone ringing etc.), unusual details (like hidden tattoo), superfluous details (not relevant to investigation; e.g. allergic to cats), related external associations (e.g. perpetrators talked about similar cases), and accounts of subjective mental state (I felt afraid, I was thinking of how to escape). CBCA has been used widely in German courts and is permissible in some courts in the USA. But it is used mainly for data gathering. Although the average accuracy rate is reportedly around 72-76%, of the various criteria, only two appear to

have produced consistent results in both field and laboratory studies, these are quantity of details, and contextual embedding (Vrij, 2005).

Significantly, these criteria also form part of a relatively more recent approach to the analysis of verbal cues from statements known as Reality Monitoring (RM) (Masip, Sporer, Garrido & Herrero, 2005; Vrij, 2005). The idea behind RM is that, because they are lying, liars have less knowledge and experience of what they are actually reporting, so truth tellers will report more: contextual information (putting information in a context, such as, 'I collected my car from the car park at 5.00 because it closes at 5.15'), and sensorial detail (including visual details, auditory details, details of emotions, and details of taste, smell and touch). In contrast, because they are internally generating information, RM hypothesises that liars will show more cognitive operations. So liars will report more thoughts and reasonings ('I must have had my coat on it was very cold that night'), and suppositions of sensory experiences ('she seemed quite clever'). Again, work using RM criteria has shown it can successfully discriminate between truthful and dishonest statements, and may outperform CBCA in this respect (Masip et al., 2005; Vrij, 2005). A major disadvantage of CBCA and RM, however, is that, to be successfully applied, the sender needs to make a fairly extensive statement which refers to an actual or fictitious event. Often in forensic investigations this kind of extensive information is not available; for example, appeals concerning missing or murdered relatives frequently contain little if any reference to sequences of events; also the sender may deny having any knowledge of an event.

Other studies have identified a number of further verbal and paraverbal behaviours that appear to relate to deception across high stakes situations. For example, speech dysfluency and errors were higher in deceptive statements of

criminal suspects made to assistant district attorneys (Davis et al., 2005), a finding which the authors attributed to anxiety. They were also higher in deceptive statements made in the police interview of a murder suspect (Vrij & Mann, 2001a), a finding which the authors attributed to cognitive load. Pausing increased in deceptive statements made by suspects in police interviews (Mann, Vrij & Bull, 2002), and again in the police interview of a murder suspect (Vrij & Mann, 2001a), findings which in both cases the authors suggest are related to cognitive load. Word and/or phrase repetition was related to deception in the studies already mentioned by Davis et al. (2005) and by Vrij and Mann (2001a), and also in a study by Harpster, Adams and Jarvis (2009), which analysed 911 homicide calls for indicators of guilt in the caller. The use of equivocal or tentative language, which Zuckerman et al. (1981) suggested liars may use in an attempt to distance themselves from their own communication, was found to increase in deceptive written statements from suspects/witnesses in police investigations (Adams & Jarvis, 2006), and in deceptive pleas for help to find missing or murdered relatives (ten Brinke & Porter, 2012). High vocal pitch was related to deception in suspect interviews when measured with specialist equipment (Hall, 1986), and in deceptive appeals for missing or murdered relatives using observer ratings (Wright Whelan, 2009), and this is in line with meta-analytic findings on vocal pitch from mostly low stakes studies (DePaulo et al., 2003).

Findings relating to other verbal and paraverbal cues have been less consistent. For example, the use of filled pauses (such as, ‘ah’ and ‘mm’), when measured separately from general speech disturbances, was negatively related to deception in Davis et al.’s (2005) study, and also in a further study of the deceptive speech of a convicted murderer (Villar, Arciuli & Mallard, 2011), but unrelated to

deception in Vrij & Mann's study (2001a). Vrij and Mann report that the murder suspect in their study produced a slower speech rate in the deceptive sections of the interview, however, Davis et al. (2005) report no difference in speech rate between honest and deceptive communications in their sample.

It appears, then, that some verbal and paraverbal behaviours may be useful in discriminating between deception and honesty in some high stakes situations, although the relevance of some behaviours across high stakes situations needs further investigation.

### ***2.2.2. Nonverbal Cues to Deception***

With regard to nonverbal cues, a research project that investigated beliefs about deceptive behaviour in 75 countries, concluded that the belief that liars avert their gaze is a dominant pan-cultural stereotype (The Global Deception Research Team, 2006). Research findings have, however, been mixed. Hence Mann et al. (2002), in an analysis of the behaviours of suspects in police interviews, found no significant effect of veracity on gaze aversion. Indeed, because of the lack of experimental support, some have dismissed gaze aversion as a cue to deception as a popular myth (for example, Vrij, 2004). In contrast, DePaulo et al. (2003) report in their meta-analysis that gaze aversion emerged as a cue to deception when motivation to succeed in the lie was a moderating variable. Furthermore, in Vrij & Mann's (2001a) analysis of the behaviour of a murder suspect, they report that he displayed more gaze aversion when being deceptive in the section of the interview before he confessed, (i.e. when he still had high motivation to succeed in the lie). Moreover,

Wright Whelan (2009) found that individuals lying when making appeals about missing and murdered relatives were perceived as producing more gaze aversion than honest appealers. The reason for the disparity in findings on gaze aversion is unclear. As noted previously, gaze aversion has been related to cognitive load (Vrij & Mann, 2001a), emotional responses such as shame (DePaulo et al., 2003), and guilt (Ekman, 1992), and attempts by liars to be evasive and distance themselves from their communications (Zuckerman et al., 1981); all factors which might be more prevalent in high stakes situations. Nevertheless, there is currently insufficient data from real life, high stakes contexts, to draw firm conclusions about the relevance of gaze aversion to deception, and more research on this topic is clearly needed.

Findings regarding other nonverbal behaviours are equally inconsistent. For instance, Koper and Sahlman (1991), in a study investigating naturally occurring, high motivation deceptive communication, report that deception was associated with a decrease in the use of self-adaptors and illustrators, which is in line with DePaulo et al.'s (2003) meta-analytic finding that deception is related to a reduction in illustrators. However, in the study previously mentioned by Davis et al. (2005), the relationship appears to change direction, as deception was related to increased gesticulation. Furthermore, Mann et al. (2002), in their analysis of the behaviours of suspects in police interviews, found no significant effect of veracity on hand movements. Similarly, although Mann et al. (2002) report that blink rate decreased during deceptive statements (a finding which the authors attributed to increased cognitive load), ten Brinke and Porter (2012) report a trend for an increased blink rate in deceptive appeals for missing and murdered relatives, a finding which the authors suggest is indicative of increased arousal.

Two behaviours which appear to have more consistency across high stakes situations are falsifying facial expressions, and headshaking. Both Davis et al. (2005), and ten Brinke and Porter (2012), report that deception was associated with attempts to produce facial expressions that were not genuine. Davis et al. (2005) also report that deceptive statements were associated with protracted headshaking, and Mann, Vrij and Bull (2006) report that suspects in police interviews produced more headshakes when lying.

As with verbal and paraverbal behaviours, it would seem that some nonverbal cues to deception may vary across high stakes contexts, even changing direction as a predictor in the case of illustrators.

### ***2.2.3. Subjective Measures as Cues to Deception***

As noted in Chapter 1, DePaulo et al. (2003), in their meta-analysis, suggest that honesty and deception may be discriminated more powerfully by using subjective measures rather than objective measures, and yet there are, to the author's knowledge, only four studies that have investigated subjective measures as predictors of deception in real life, high stakes contexts. Koper and Sahlman (1991), in their investigation of public figures making televised statements, found that deception was related to increased perceived arousal, tension, and vocal tension, and also that liars were perceived as being more involved, immediate and interested in their presentation. Truth tellers were perceived as being more plausible and consistent than liars, and to have more certain and direct voices. Wright Whelan (2009), in a study of appeals for missing and murdered relatives, also reports a trend for honest

appealers to be perceived as being vocally more direct than deceptive appealers. Some of these findings are in line with DePaulo et al.'s meta-analytic findings of mostly low stakes studies, in which deception was related to increased discrepancy, verbal and vocal uncertainty, tension, and vocal tension, whereas honesty was related to increases in plausibility and logical structure. However, DePaulo et al. (2003) report that verbal and vocal involvement, and immediacy, were related to honesty, whereas Koper and Sahlman (1991) report that these behaviours were related to deception. Other results also offer further contradiction; Mann and Vrij's (2006) study of suspect police interviews found that suspects were perceived as being more tense when making honest statements than when making deceptive statements, whereas Vrij and Mann's (2001a) study of a murder suspect's police interview found that he was perceived as more tense when being deceptive than when being honest. Both studies found that deception was related to subjective impressions of the interviewees thinking hard, but whereas one study (Mann & Vrij, 2006) found deception to be related to subjective impressions of the interviewees controlling their behaviour, the study of the murder suspect found no such effect (Vrij & Mann, 2001a).

Again, as with verbal and paraverbal behaviours, and nonverbal behaviours, although there appear to be some subjective measures which may be useful cues to deception across high stakes contexts, there are clearly discrepancies in the findings and this area also warrants further investigation. In accounting for disparities between studies, one particular area that has clearly been neglected is the possible role of context, an issue which is discussed below.

### **2.3 The Importance of Context**

The issue of context has not been fully addressed in the literature on deception, and the contradictory findings in existing high stakes deception research may, in part, be due to a lack of contextual focus. For example, DePaulo et al. (2003) have suggested that some cues may occur more often, or even only, in forensic contexts. However, it is possible that some cues may be even more context-specific than this.

A number of researchers have suggested that a range of situational factors may affect the production and salience of cues related to deception (see, for example, DePaulo & Morris, 2004; Porter & ten Brinke, 2010; Sporer & Schwandt, 2006); thus different processes related to the production of behaviours related to deception may be more relevant in different situations, so that, for example, nervous responses are likely to be most salient in very high stakes situations, whereas cues related to cognitive load are more likely to be more salient in complicated lies (Vrij & Mann, 2004). However, studies investigating high stakes lies often lack a contextual focus; for example, Mann et al.'s (2002) study of police interviews included suspects who were being investigated for crimes as diverse as burglary, rape, and murder; Koper and Sahlman's (1991) study investigating naturally occurring, high motivation deceptive communication, included athletes, politicians, and military personnel, telling a variety of types of lie; and Davis et al.'s (2005) study of statements of criminal suspects made to assistant district attorneys, included crimes from murder, to sexual assault, and possession of illegal weapons.

As it is highly likely that different types of lies may have different emotional, motivational and cognitive correlates, different types of lies may not necessarily resemble each other in terms of behavioural responses. For example, a career



criminal lying about a burglary in a police interview may not experience the same emotions (i.e. shame, guilt, and fear) as somebody lying about the spontaneous murder of a close relative; or somebody falsely confessing to a crime may not experience the same motivational drive as somebody committing a high value insurance fraud. This proposition is in line with findings by O'Sullivan and Ekman (2004), that different groups of people are more successful at identifying types of lie with which they are most familiar. In their research, they found that law professionals were significantly more successful at identifying lies in a crime deception task than in an emotional deception task, whereas therapists showed the opposite pattern. This suggests that all lies do not necessarily resemble each other in terms of the responses they evoke.

In a study directly investigating this possibility, Wright Whelan (2009) found that cues to deception were not comparable across two high stakes, forensic contexts: public appeals for help with missing or murdered relatives, and confessions. Deceptive appealers (i.e. those involved in the murder of their relative) were perceived to have significantly higher vocal pitch and more speech hesitations than honest appealers, and large (non-significant) effect sizes in the expected direction for several further cues, whereas people making false confessions produced none of the cues to deception investigated in the study. ten Brinke and Porter (2012), in their analysis of appeals for help with missing or murdered relatives, note that several cues related to affective responses discriminated between honest and deceptive appeals, and only one cue related to cognitive load, a finding which the authors suggest supports the proposal that in certain contexts, behaviours related to the affective experience of the liar may be more salient than behaviours related to increased cognitive load.

Furthermore, as noted above, although a study of police suspect interviews found a decrease in blink rate to be associated with deception (Mann et al., 2002), ten Brinke and Porter (2012) report a trend for an increased blink rate related to deception, a finding which the authors suggest is associated with arousal, and which counters the notion that cognitive load alone can account for changes in blink rate during deception. Indeed, although undoubtedly an extremely high stakes context, appeals for help with a missing or murdered relative often do not include an account of an event, and, therefore, at the time, deceptive appealers may not be involved in the cognitive task of ‘keeping their story straight’.

This lack of contextual focus in studying high stakes lies may not only ‘muddy the waters’ regarding which cues discriminate in which contexts, it may also prevent previously uninvestigated, context-specific cues, from emerging. As noted earlier, DePaulo et al. (2003) suggest that it may be important to establish a set of cues that are relevant within a particular context, even if the cues cannot be generalised to different contexts. Two studies using a tightly-focused, high stakes context, have yielded interesting results in this regard. In Harpster et al.’s (2009) study of 911 homicide calls, 19 behaviours were found to discriminate between innocent callers and callers who were later found to be involved in the homicides they were reporting. Importantly, a number of these behaviours had not been previously identified and were likely to be specific to the particular context; for example, acceptance of death, insulting the victim, inappropriate politeness, and location in the call of the plea for help. Furthermore, ten Brinke and Porter (2012), in their analysis of the behaviours of people appealing for help with missing or murdered relatives, found facial expressions related to disgust, distress, surprise and happiness discriminated between honest and deceptive appealers, again, behaviours

that are likely to apply only in specific contexts. These findings suggest that a more contextual focus might be beneficial when examining high stakes lies.

The findings reported by Harpster et al. (2009), and by ten Brinke and Porter (2012), may also have another important implication. As the dominant theoretical approaches to the production of cues to deception outlined earlier (cognitive load, affective responses, behavioural control and impression management, and arousal), do not fully account for some of the context-specific cues identified in these two studies, it may be that new theoretical approaches are required to better accommodate the more context-specific cues that may emerge only in high stakes situations. For example, in Harpster et al.'s study of 911 homicide calls (2009), the cues acceptance of death, insulting the victim, inappropriate politeness, and location in the call of the plea for help, are not readily explained by the Four Factor Model (Zuckerman et. al, 1981). The leakage of emotion in facial expression, and the failure to produce genuine facial expressions identified by ten Brinke and Porter (2012), may be accounted for by the affective response factor, but this general approach does not specifically predict the emotions which are reported in the study; the particular emotions (disgust, distress, surprise and happiness) are likely to be relevant as cues to deception only in specific contexts. A theoretical approach which accommodates these high stakes, context-specific cues is lacking in the literature, hence this is another issue which the present thesis attempts to address.

## **2.4 Conclusions**

To summarize, it obviously seems important to study deception in real life, high stakes situations, as well as in the laboratory; as, although some findings regarding cues to deception may occur in both types of context, some cues may only emerge, or even change direction, in high stakes situations. However, the issue is made more complex by the fact that, even in investigations of real life, high stakes situations, some results achieved so far have not been consistent across studies. This is perhaps not surprising; the body of literature on high stakes lies is as yet not sufficiently developed to produce robust findings. In general the sample sizes of deceptive and honest communications have been small; indeed, in some cases, the data have been produced by just one person (Villar et al., 2011; Vrij & Mann, 2001a). Consequently, one of the overarching aims of the present thesis was to add to this literature by further investigating nonverbal, verbal, and subjective cues to deception, in a real life, high stakes context.

Also, the nascent body of research investigating cues to deception in real life, high stakes situations, suggests that context may affect the production of some cues to deception, and furthermore, previously uninvestigated, context-specific cues may exist. Those few studies that have utilised a tightly-focussed, context-specific approach, have produced findings which are not readily accommodated by existing theories of the production of cues to deception. There is, therefore, a marked lacuna in the literature for research with a clear focus on deception in specific contexts in real life, high-stakes situations. Consequently, the present thesis focused on the specific, high stakes context of public appeals for help with missing or murdered relatives.

# **Chapter 3**

## **Accuracy in Detecting Deception in High Stakes Contexts**

Chapters One and Two have highlighted the role that ecological validity may have on the production of cues to deception. In particular, it has been argued that situations involving high stakes lies are more likely to produce salient, identifiable cues to deception, than low stakes situations. As the detection of deception is obviously dependent on the cues available, it follows that people are going to be more able to accurately detect lies in situations in which the cues are most numerous and prominent; i.e. high stakes situations. However, only a few studies have examined observer accuracy in the detection of real life, high stakes lies. These are discussed below.

### **3.1 The Effect of Ecological Validity on Accuracy**

As previously noted, meta-analytic findings suggest that observer accuracy in detecting deception, at 54%, is barely above chance (Bond & DePaulo, 2006). One possible explanation for the generally low accuracy rates, is that they are an artefact of the stimulus materials used in the experiments; i.e. the large majority of studies of

deception detection have used low stakes lies as stimulus materials, in which cues to deception may be faint and unreliable.

The few studies that have investigated observer accuracy in detecting deception in high stakes situations, have produced more promising, and reasonably consistent, results. In one study (Vrij & Mann, 2001b), using appeals for help with missing or murdered relatives as stimulus materials, police officers were unable to distinguish honesty from deception above the level of chance. However, this study was limited in that accuracy results were based on only five appeals, all of which were deceptive, and were not in the native language of the participants. In contrast, in a series of studies using videos of police interviews with suspects as stimulus materials (Mann & Vrij, 2006; Mann, Vrij & Bull, 2004; Mann et al., 2006; Vrij & Mann, 2001a; Vrij, Mann, Robbins & Robinson, 2006), police officers' credibility judgements were consistently accurate at rates well above chance levels, at 68%, 65%, 69%, 64% and 72% respectively.

It can be noted that, within these studies, there were some variations in the occurrence of a truth bias. As mentioned in Chapter 1, meta-analytic reports of predominantly low stakes studies indicate that observers are more likely to judge a communication as honest than as deceptive, and thus truthful messages are correctly identified more frequently than deceptive messages (Bond & DePaulo, 2006). The findings regarding a truth bias in the small body of research utilising high stakes lies is mixed: in three of the studies no truth bias was found (Mann & Vrij, 2006; Mann et al., 2004; Vrij et al., 2006), in one study a significant truth bias was found (Vrij & Mann, 2001a), and in one study again more truths (73%) were correctly identified than lies (64%), although inferential statistics were not reported (Mann et al., 2006).

Nevertheless, the generally high accuracy rates reported in the above studies are important, as previous studies, using low stakes stimulus materials, have suggested that police officers are no more accurate than laypersons at detecting deception. For example, DePaulo and Pfeifer (1986) found no differences in accuracy between students (54%), new recruits to a federal law enforcement training programme (53%), and experienced federal law enforcement officers (52%), when judging people lying or telling the truth about their opinions. Similarly, Ekman and O'Sullivan (1991) report that police officers (56%) were not significantly more accurate than students (53%) when viewing participants lying or telling the truth about their response to a film that they were watching. And Meissner and Kassin (2002) used videos of students being interviewed lying or telling the truth about mock crimes, and found that police investigators (50%) were not more accurate than students (56%). Vrij (1993) also reports that detectives achieved an accuracy rate of 49% when judging participants lying about whether or not they were in possession of a set of headphones.

It would appear then, that police officers' ability to accurately detect deception increases when they view police interviews of suspects. However, due to restrictions placed on the viewing of the videos of the interviews, the authors of the aforementioned studies were unable to test whether non-police participants would also achieve higher accuracy than is usual in deception detection studies, leaving a number of important questions unanswered. For example, the higher accuracy rates may have been due simply to the use of high stakes lies, which may have magnified the cues used to differentiate between honest and deceptive behaviour, and so made them more easily discernible by all. Alternatively, it could be that police officers are particularly good at detecting deception in police interviews, due to their familiarity

with this context. As noted in Chapter 2, due to a domain familiarity effect, different groups of people may be more successful at identifying types of lies with which they are most familiar (O'Sullivan & Ekman, 2004). However, we do not know whether police officers would achieve comparable high accuracy in distinguishing truths from lies in other high stakes contexts.

### **3.2. Accuracy and Confidence in Veracity Judgements of High Stakes Lies**

As mentioned previously, meta-analytic reports indicate that there is no relationship between accuracy and confidence in making judgements of veracity (DePaulo et al., 1997), so that observers are no more confident when making correct judgements than when making incorrect judgements of deception. Furthermore, several studies report that police officers in particular tend to be more confident in their credibility judgements, than laypersons. For example, both DePaulo and Pfeifer (1986), and Meissner and Kassin (2002), report that police officers were more confident in their veracity judgements than students, despite being no more accurate, and Vrij (1993) reports that detectives were very confident in their veracity judgements despite an average accuracy rate of 49%.

However, again, these studies involved observer judgements of low stakes lies, and the few results from studies utilising high stakes lies as stimulus materials have been mixed. In two studies in which police officers viewed police interviews of suspects lying and telling the truth, no relationship between confidence and accuracy was found (Mann et al., 2004; Vrij & Mann, 2001a). However, another study of police officers viewing police interviews of suspects, found that participants were



more confident in their correct judgements than in their incorrect judgements (Mann et al., 2006). Again, due to restrictions placed on the viewing of the videos of police interviews with suspects, the authors of this body of research were unable to investigate the effects of high stakes stimulus materials on the accuracy-confidence relationship of non-police participants.

Clearly, further research is needed to clarify what effect, if any, the use of high stakes stimulus materials has on observer confidence, and the relationship of this to observer accuracy in detecting deception. It may be that if observers are basing their credibility judgements on more pronounced cues to deception, they may be more accurate and also more confident in their accurate decisions.

### **3.3 Individual Variation in Accuracy in Detecting Deception**

It is usual in deception detection studies for accuracy rates to be reported for the group of observers as a whole; however, in a handful of studies, researchers have reported individual differences in accuracy between observers, and large differences have emerged among them.

For example, in a study in which both police and student participants viewed videos of people lying or telling the truth about whether or not they were in possession of a set of headphones, Vrij and Graham (1997) report that individual accuracy rates ranged from 20% to 90%. Similarly, in a study in which students and law enforcement personnel viewed videos of paroled felons making various honest and deceptive statements (for the sake of the experiment), Bond (2008) reports identifying two individuals, both correctional officers, who scored above 80%

accuracy on both parts of a repeated deception detection task. Three studies investigating police accuracy when viewing high stakes lies produced similar results. In a study in which 52 police officers viewed deceptive people appealing for help with missing or murdered relatives, Vrij and Mann (2001b) report that three officers achieved a high accuracy score of 80%, whereas three officers achieved accuracy scores of 20% and one officer scored 0%. Similarly, in a study in which 65 police officers viewed truthful and deceptive fragments from a police interview with a murder suspect, Vrij and Mann (2001a) found that six officers achieved 100% accuracy, and five officers achieved a high accuracy score of 83%, whereas 11 officers achieved lower than 33% accuracy. Finally, in a study in which police officers viewed police suspect interviews, Mann et al. (2006) report identifying four officers who achieved 90% accuracy. It should be noted that the studies mentioned above were all limited by the use of a relatively low number of 'senders' (i.e. people producing the honest/deceptive communications); from one suspect producing six fragments (three honest, three deceptive), to 14 suspects (each producing honest and deceptive fragments).

It seems, therefore, that although there has been a lively debate regarding the existence of experts in deception detection (see, for example, Bond & Uysal, 2007; O'Sullivan, 2007), the evidence suggests that substantial individual variations in the ability to detect deception do exist. This, perhaps, should not be surprising. To detect deception, two processes must occur; first, there must be some differences in behaviours between honest and deceptive communications; and second, the observer must perceive these differences and understand them. As has been discussed, research indicates that there are some differences between honest and deceptive behaviours, even if these differences tend to be small and may not be consistent

across all contexts; hence it might be expected that some people are better able than others to perceive these differences and to understand them.

Consequently, some attempts have been made to investigate factors, both innate and acquired, which may be related to individual differences in the ability to detect deception; for example, higher accuracy in deception detection has been variously related to left-handedness (Porter et al., 2002), a greater Need for Cognition (Reinhard, 2010), and the personality factors Openness to Experience, and Agreeableness (Enos et al., 2006). O'Sullivan and Ekman (2004) began a programme to investigate characteristics of accurate lie detectors that they had previously identified; in their initial report they comment that their 'wizards' of deception detection share a sensitivity to the nuances of language use, unusual childhoods, and high motivation. In another study, Mann et al. (2004) found that accuracy was significantly correlated with interviewing experience in a sample of police officers; however, other studies (Vrij & Graham, 1997; Vrij & Mann, 2001a; Vrij & Mann, 2001b) have found no relationship between accuracy and interviewing experience.

Although a detailed investigation of the variables that may affect accuracy in deception detection is beyond the scope of the present thesis, an important question that can and is addressed is, if some individuals are able to accurately detect deception, what are the cues on which they base their credibility judgements? It seems likely that the behaviours on which they base their judgements are reliable, and yet there has been very little research investigating this.

The method of investigating cues used by accurate lie detectors is nascent. For example, Mann et al. (2006) asked police officers with high deception detection accuracy rates to select fragments from clips of suspect interviews that they found most useful, and then compared accuracy ratings between subjects viewing whole

clips, with subjects viewing just the fragments. Subjects were not more accurate in detecting deception when viewing only the fragments selected by accurate people, but they had not been told specifically which behaviours were considered important by the accurate participants. O'Sullivan and Ekman (2004) have also begun to investigate how their 'wizards' are able to accurately detect deception, and a few preliminary findings have been published: the 'wizards' report a frequent use of nonverbal behaviour, and use of inconsistencies and consistencies between verbal and nonverbal behaviour, as well as demonstrating high sensitivity to the nuances of language use. Bond (2008) found that accurate deception detectors reported using more nonverbal cues, and fewer verbal cues, than a control group. Also whilst one of the accurate deception detectors focused more on the face, the other focussed more on the arm/torso area.

Clearly then, there is substantial scope for further research in this area. As previously noted, DePaulo and Morris (2004) suggest that there may be cues to deception other than those which have been identified so far in laboratory research, which have yet to be investigated; an inductive, bottom-up approach, exploring the cues used by accurate deception detectors may be a useful method for generating previously unidentified cues, as well as expanding our understanding of the complexities of deceptive behaviour.

### **3.4. Consensus Judgements**

An issue that has not yet been addressed in the deception literature is the value of consensus judgements as predictors of veracity. If there is a strong consensus

amongst observers that a ‘sender’ (i.e. person producing the honest/deceptive communication) is lying, or telling the truth, how likely is it that the consensus judgement is correct? It could be that different observers may have different abilities to detect different types of cues and that, by combining these abilities, they might collectively be more accurate. In a meta-analysis of mostly low stakes studies, investigating sender and observer variance on deception detection accuracy, Bond & DePaulo (2008) found that although there was large variance in sender demeanour, there was little variance in observer ability in making accurate credibility judgements.

However, this issue has not yet been investigated using real life, high stakes stimulus materials. As discussed earlier, high stakes situations may elicit more numerous and more discernible, cues to deception, and it is not known whether these factors will affect consensus judgements, and whether consensus judgements have any utility in deception detection.

### **3.5 Conclusions**

To summarize, it would appear that ecological validity may potentially have a substantial effect on the accuracy of deception detection, and may also affect truth bias; however, so far research on this has been limited to only police observations of suspect interviews. It is, therefore, unknown whether the effects of high stakes stimulus materials on accuracy in detecting deception generalise to non-police populations, and to other high stakes contexts. Consequently, the present thesis

investigated accuracy in detecting deception in police and non-police participants, viewing a real life, high stakes situation.

In addition, relatively little is known about the relationship between accuracy and confidence in judgements of veracity in high stakes situations. Published findings have been mixed, and studies have again been limited to police participants, and only one context (police interviews of suspects). To the author's knowledge, there is no published research investigating the relationship between accuracy and confidence in judgements of high stakes deception in the general public. Consequently, the present thesis also investigated the accuracy-confidence relationship in veracity judgements of police and non-police participants, when viewing high stakes lies.

Individual variation in the accuracy of credibility judgements is another area that has received little attention in the deception literature. A few studies have made some limited attempts to explore this area, however, it may be useful to investigate more fully the cues on which those who are better able than most to detect deception, base their judgements. Hence this approach was adopted in the present thesis.

Finally, as far as the author is aware, no published research has investigated whether observer consensus on the credibility of a sender is a good predictor of veracity in a high stakes context. Consequently, this was also explored in the present thesis.

## **Part II**

### **The Empirical Research**

# **Chapter 4**

## **Introduction to the Empirical Research**

As an introduction to the empirical section of the thesis, it may be beneficial to briefly summarise the main points introduced in Part I regarding the rationale for, and aims of, the research presented here. Based on a review of the existing literature, the central focus of this thesis was to explore cues to deception in a real life, high stakes, situation, and also to investigate observer accuracy in detecting deception when viewing real life, high stakes lies. It was, thereby, hoped that the present research would contribute to the small body of existing research on high stakes deception in several key ways; by reinforcing previous findings regarding cues to deception which may be useful across high stakes contexts; by identifying previously uninvestigated, and possibly context-specific, cues to deception; by using a multiple-cue approach to behaviours associated with deception in a specific high stakes context; and to investigate police and non-police accuracy in detecting deception in a high stakes situation.



#### **4.1. Investigating Naturally Occurring, High Stakes Deception**

As discussed in Chapter 1, ecological validity has been identified as a key issue by deception researchers (for example, DePaulo & Morris, 2004; Frank & Svetieva, 2012; Granhag & Stromwall, 2004; Porter & ten Brinke, 2008; Shuy, 1998; Vrij, 2004), who agree that there are likely to be differences between cues exhibited in low stakes situations and those exhibited in high stakes situations. The large majority of existing deception research has utilised low stakes lies, and while the results of these low stakes studies may be relevant to low stakes situations, the relevance of their findings to high stakes contexts needs further investigation. The presence of high stakes may exacerbate various factors which are relevant to the production of cues to deception (including cognitive load, affective responses, behavioural control, and arousal), potentially resulting in more numerous and more prominent cues to deception. Furthermore, the use of high stakes lies as stimulus materials may have an effect on observer accuracy; if cues to deception are more marked in high stakes contexts, observers may be more able to detect deception than in low stakes contexts.

It is in high stakes contexts that the consequences of deception are most serious, and yet there have been only a handful of studies investigating naturally occurring, high stakes deception. As Fuller, Biros, Burgoon and Nunamaker (2013) have pointed out, there are a number of factors which have limited research using high stakes field data, one of the most significant being that ground truth must often be surmised rather than verifiable with complete certainty. This is indeed a challenge, and yet if it can be reasonably overcome, there are important advantages in using data that are ecologically valid. Indeed, after symposium in 2003 on deception, the organisers argued that, ‘Deception researchers will need to step out of

the laboratory and into real-life forensic contexts to find the most wanted answers' (Granhag & Stromwall, 2004, p.324).

To address this issue, therefore, the research in the present thesis needed stimulus materials that were derived from real-life situations. The materials that were chosen consisted of video footage of people making public appeals for help with missing or murdered relatives. It is not unusual, when a person goes missing or is killed, for a close relative to appear before the press and appeal to the public to help find the missing person, or to help find out who killed the person. The relative making the appeal is usually a parent or step-parent, a spouse or partner, a son or daughter, or a sibling. Sometimes the person making the appeal is honest; he or she is not involved in the death or disappearance of the relative, and is genuinely appealing to the public for help. However, sometimes the person making the appeal is deceptive; he or she is involved in the death or disappearance of the relative, and the appeal is a deceptive communication, in which the appealer attempts to manipulate the beliefs of others by concealing knowledge of the crime, and falsifying an appeal (a deceptive appealer does not actually want the public to help find the relative, or find out who killed the relative). The rationale for the choice to use appeals as stimulus materials is as follows.

#### **4.2. Investigating the Specific Context of Public Appeals**

As discussed in Chapter 2, the issue of context has not been fully addressed in the literature on deception. Clearly, it is important to establish cues that have utility in discriminating truth from deception across high stakes situations, and previous

research in this area suggests that there are some cues to deception which may be useful in this way. However, a range of situational factors may affect the production and salience of cues to deception, so that even in high stakes situations, lies in different contexts may not necessarily resemble each other (see, for example, DePaulo & Morris, 2004; Porter & ten Brinke, 2010; Sporer & Schwandt, 2006). For example, in the emotionally charged context of making an appeal, it could be expected that cues related to affective responses would be particularly salient, and as an appeal often does not contain an account of an event, cues related to cognitive load may be less prominent. Research using a tightly-focused, specified context, may help clarify some of the mixed findings in previous studies investigating high stakes deception, and produce more consistent findings. DePaulo et al. (2003) have pointed to the importance of establishing a set of cues that are relevant within a particular context, even if the cues cannot be generalised to different contexts, and it is possible that investigating a specific context may allow previously unidentified cues to emerge.

Notably, the two studies using a high stakes situation which focused on a tightly-defined context have produced findings which suggest that context-specific cues may exist in certain high stakes situations (Harpster et al., 2009; ten Brinke & Porter, 2012); however, studies investigating high stakes lies have generally lacked a contextual focus. The situation of public appeals for help with missing or murdered relatives obviously has a strong contextual focus, which was a major reason for choosing it. However, this context also has a number of other advantages.

First, as mentioned in Chapter 1, in their meta-analysis of cues to deception, DePaulo et al. (2003) found that strong motivation to succeed in the lie, lies which were identity relevant, and lies about transgressions, were all moderating factors that

produced more prominent cues to deception. Appeals could be considered a context which is very likely to include these three factors. DePaulo et al. (2003) define a transgression as a behaviour regarded as serious and which elicits punishment or disapproval, and being involved in the death or disappearance of a relative clearly fits with this definition. Deceptive appealers could be expected to have strong motivation to succeed in their deception, as the punitive consequences of failing include long prison terms or even, in some places, capital punishment. Also, lying about being involved in the death or disappearance of a relative could be expected to be 'identity relevant' as defined by DePaulo et al. (2003); that is, concealing a discreditable deed could be construed as protecting the liar's image.

Second, an appeal presents an opportunity for the researcher to examine behaviours that are untainted by interviewer questions or demeanour. A large amount of the existing research investigating behaviour related to high stakes deception has involved police interviews of suspects (for example, Mann & Vrij, 2006; Mann et al., 2002, 2006; Vrij & Mann, 2001a); this may be problematic as an interview is a form of two-way communication, and suspect behaviour is likely to be influenced by the interviewer (van Koppen, 2012). It should be acknowledged that in a small minority of appeals, the appealers are asked questions by reporters, but in general, an appeal is an uninterrupted communication.

Third, one consequence of the use of police suspect interviews in a large part of the existing body of research on high stakes deception (Mann & Vrij, 2006; Mann et al., 2004, 2006; Vrij & Mann, 2001a; Vrij et al., 2006), is that viewing of the stimulus materials has been restricted to police officers. The use of appeals allowed non-police participants to view the stimulus materials, and to compare non-police and police accuracy in detecting deception in a high stakes situation, and also

allowed the investigation of police accuracy in detecting deception in a context other than a police interview.

Fourth, although some have stressed the importance of using active interactive situations to study cues to deception (DePaulo & Bond, 2012; Vrij & Granhag, 2012), the method in which observers make judgements of deception by viewing short video clips in a passive paradigm reflects exactly how appeals are investigated in real life. Appeals are watched passively by observers, including police officers who are investigating the case, and who may well be forming credibility judgements about the appellant during the appeal. It is important, then, that there exist some criteria to guide these credibility judgements.

Finally, on a practical level, appeals are a high stakes situation of which video footage is readily available. Viewing video of high stakes situations in a forensic domain is often restricted and difficult to access, whereas videos of appeals are already in the public domain and therefore more easily accessible. Using video (instead of, for example, written statements) also allowed the investigation of verbal, paraverbal, and nonverbal behaviours, allowing a multiple-cue approach across channels of behaviour. Such materials also allowed access to information that could be used to establish ground truth concerning guilt and innocence. This issue will be explored in more detail later.

#### **4.3. Research Aims and Questions Relating to the Identification of Cues to Deception**

In the light of the issues discussed so far, and as identified in the Preface, the thesis had four major aims; three relating to the identification of cues to deception, and one to observational accuracy. The three major aims relating to the identification of cues to deception were as follows.

1. To develop and apply a methodology to identify cues to deception in a specific ecologically valid high stakes context, that of appeals for help with missing and murdered relatives, and to investigate whether these cues are useful in discriminating between honest and deceptive appeals in this context. It was hoped that this methodology would suggest previously unidentified cues, and context-specific cues, for further investigation.
2. To develop a multiple-cue approach to behaviours which discriminates between honesty and deception in the context of appeals, and which may be useful in an applied context.
3. To develop a theoretical rationale for cues that may be context-specific and may emerge only in high stakes contexts, and cannot readily be accommodated by existing theoretical approaches.

Within these general aims, there were several additional, more specific research questions and hypotheses, based on the literature review as follows.

- Are there cues to deception identified in research on high stakes lies in other contexts, which also discriminate between truth and deception in the high stakes context of appeals? It was hypothesised that some cues would be related to deception across high stakes contexts.
- Are there previously unidentified cues to deception which may be specific to the context of appeals? It was hypothesised that some new cues to deception would emerge from the research.
- Can subjective measures be used to discriminate between honesty and deception in the context of appeals? It was hypothesised that subjective observer ratings would produce some discriminatory cues.
- In the emotionally charged context of appeals, are cues related to affective responses particularly salient? It was expected that cues related to cognitive load would be less prominent than cues related to emotion in this context.
- Does the high stakes context of appeals exacerbate factors which underlie the production of cues to deception and magnify them, so that effect sizes are larger than is usual in low stakes deception research? It was hypothesised that effect sizes for cues found to discriminate between honest and deceptive appeals would be substantial.
- Experimental, laboratory-based studies, in which participants are randomly allocated to groups, do not allow for possible individual differences to emerge in the type of person who may choose to deceive, particularly in high stakes situations. Do some cues emerge which may be associated with individual differences, rather than acts of deception per se?

#### **4.4 Research Aim and Questions Relating to Accuracy in Detecting Deception**

The major aim relating to accuracy in detecting deception was as follows.

4. To investigate the effect of using ecologically valid stimulus materials on the accuracy of observers (both police and non-police) in detecting deception, on truth bias, on accuracy-confidence relationships, and the utility of consensus judgements as a predictor of veracity.

Again, within this general aim, there were several additional, more specific research questions and hypotheses based on the literature review as follows.

- Does the use of ecologically valid stimulus materials affect observer accuracy in detecting deception? It was hypothesised that the use of real life, high stakes lies, would result in all participants (police and public) achieving above chance accuracy in detecting deception
- Are police observers more accurate than non-police observers when making veracity judgements about real life, high stakes lies? It was hypothesised that, due to the likelihood of increased exposure to deceptive behaviour in the course of their jobs, police officers would be more accurate at detecting real life, high stakes deception, than the general public.
- Are police observers accurate when viewing a real life, high stakes context other than a police interview? It was hypothesised that, because deceptive appeals involve lying about a serious crime, police officers' accuracy would be equivalent to that in previous research using police interviews.



- What is the effect of ecological validity on the existence of a truth bias for both police and non-police observers? It was hypothesised that the use of high stakes lies would eliminate truth bias. It was also expected that, again due to the nature of their jobs, police officers may show evidence of a deception bias.
- What is the effect of ecological validity on the accuracy-confidence relationship for both police and non-police observers? It was hypothesised that the use of real life, high stakes deceptive behaviour, would result in a positive relationship between accurate veracity judgements and confidence.
- Does observer consensus predict veracity in a high stakes context? Although not previously researched, it was expected that consensus would predict veracity.
- Do the most accurate observers use the cues identified by the research as discriminating between honesty and deception in this context? Again although there has been no previous research on this issue, it seemed reasonable to predict that accurate observers would rely on some of the cues identified in the present thesis.

#### **4.5 Organisation of the Studies**

With the aims and questions identified above in mind, the following studies were conducted.

#### **4.5.1. Study 1**

Study 1 was a qualitative investigation of cues to deception in appeals, reportedly used by people identified in a previous study as being unusually accurate at detecting deception in the context of appeals. The aim of the study was to generate as many cues to deception in appeals as possible, possibly new cues, and possibly context-specific cues, to be tested in future, quantitative research. Nine appeals were used as stimulus materials in this study.

#### **4.5.2. Study 2**

Study 2 was an empirical investigation of cues identified in Study 1, using subjective observer ratings, on the same set of stimulus materials as in Study 1. The aim of the study was to identify cues which were useful in discriminating between deceptive and honest appeals, to be investigated further on a larger and different sample set of appeals. There were many cues generated in Study 1, and it was necessary to filter out the cues which were worthy of further investigation. The analyses included case by case analyses.

#### **4.5.3. Study 3**

Study 3 was an empirical investigation of verbal cues amenable to frequency counts identified in Study 1, and also verbal cues suggested by previous research, on the same set of stimulus materials as Study 1. The aim of the study was to identify verbal cues which were useful in discriminating between deceptive and honest appeals, to

be investigated further on a larger and different sample set of appeals. The analyses included case by case analyses using verbal cues, and also combining cues from Studies 2 and 3.

#### **4.5.4. Study 4**

Study 4 was an empirical investigation of cues to deception in appeals, using subjective observer ratings, on a new and larger set of stimuli (32 appeals). Cues included in the study were those identified in Study 2 as being useful in discriminating between honest and deceptive appeals, cues identified in previous research on high stakes lies, and two cues noticed by the researcher. The aim of the study was to investigate whether consistent patterns of behaviour that discriminate between honest and deceptive appeals, would emerge from subjective observer ratings.

#### **4.5.5. Study 5**

Study 5 was an empirical investigation of verbal cues to deception using frequency counts, on transcripts of the new and larger set of 32 appeals, also used in Study 4. Verbal cues included in the study were those identified in Study 3 as being useful in discriminating between honest and deceptive appeals, cues identified in previous research on high stakes lies, and three cues noticed by the researcher. The aim of the study was to investigate which verbal cues discriminated between honest and deceptive appeals.

#### **4.5.6. Study 6**

Study 6 was an empirical investigation of nonverbal cues to deception using frequency counts, on the new and larger set of 32 appeals, also used in Studies 4 and 5. Cues included were three cues identified in previous research on high stakes lies; no cues specifically derived from Study 1 were included, as the only nonverbal cue amenable to frequency counts identified in Study 1 has been widely investigated in previous research. The aim of the study was to investigate which nonverbal cues amenable to frequency counts discriminated between honest and deceptive appeals. The analyses included case by case analyses, combining cues from Studies 4, 5 and 6.

#### **4.5.7. Study 7**

Finally, Study 7 was an empirical investigation of police and non-police participants' accuracy in detecting deception, when viewing the 32 appeals used in Studies 4, 5 and 6 (plus 4 'filler' appeals). The aim of the study was to investigate accuracy, confidence, truth bias, and consensus judgements, and also whether accurate participants reported using the cues identified as discriminatory in Studies 4, 5 and 6, more frequently than less accurate participants.

### **4.6. Ethical Considerations**

The studies reported in this thesis were conducted in compliance with BPS and APA guidelines and were formally approved by the University of Liverpool Committee on

Research Ethics. All participants recruited to take part in the studies were non-vulnerable adults who were able to give informed consent. Participation in the studies involved viewing videos of appeals for help with missing or murdered relatives, a type of material that is commonly shown in news programmes. Participants were provided with an information sheet before participating, informing them of the nature of the study, the type of stimulus material they would be viewing, that their participation was voluntary, and that they were free to withdraw at any time (see Appendix 1). After taking part in the studies, participants were provided with a debriefing sheet (see Appendix 2). All data collected from the participants were treated confidentially, and there was no identification of individuals in the analyses.

The stimulus materials used in the studies in this thesis were already in the public domain, and were freely available to view from various internet news and media sites. Appeals for help with missing or murdered relatives are publicly made in situations in which the person involved would expect to be observed; i.e. in press conferences or in other public environments, directly in front of news cameras. Consequently, none of the stimulus materials used in this thesis could be considered private or confidential.

# **Chapter 5**

## **Methodology: The Stimulus Videos**

As the primary method in the present thesis was to utilise real life appeals for help with missing or murdered relatives, the first step was to select a suitable sample of such materials.

### **5.1 The Stimulus Materials**

Freely available video footage of people making public appeals for help with missing or murdered relatives was collected from various online news and media sites from the United States, the United Kingdom, Canada and New Zealand. Appeals were considered for inclusion only if they satisfied the stringent criteria for establishing ground truth, discussed below. Further, appeals were not considered for inclusion if they were recent and high profile in the UK, or if they were made many months or years after the event; all appeals included were made within one month of the relative going missing or being murdered. See Appendix 3 for case summaries of all appeals used as stimulus materials.

Ten appeals (five deceptive and five honest) had been previously collected for the author's Masters dissertation (Wright Whelan, 2009), from which the two

accurate participants who took part in Study 1 of the present thesis had been identified. Hence, nine of those original ten appeals were used in studies 1, 2 and 3 (the outcome of one of the ten appeals was known to both participants in Study1, and so was not included). There were five honest appeals, two of which involved missing relatives and three of which involved dead relatives, and four deceptive appeals, two of which involved missing relatives and two of which involved dead relatives. A summary of the properties of the appeals is presented in Table 5.1 below.

Table 5.1: *Properties of appeals used as stimulus materials in studies 1, 2 and 3*

	Relative missing	Relative dead	Gender	Length in seconds <i>M (SD)</i>
<b>Deceptive appeals</b> <b>(<i>n</i> = 4)</b>	2	2	2 male 2 female	53.00 (22.96)
<b>Honest appeals</b> <b>(<i>n</i> = 5)</b>	2	3	2 male 3 female	48.80 (21.50)

An example of a deceptive appeal in this sample was a man who made a televised appeal for help in finding his wife, after she failed to arrive at work and her car was found abandoned. Her body was found shortly afterwards, and the husband was arrested after eyewitnesses reported seeing him at his wife's apparently abandoned car, the morning she disappeared. After inconsistencies were identified in the man's account, and items missing from his wife's purse were found in his shoe, he admitted to strangling his wife during a row, and dumping her body. At his trial, he pleaded guilty to second degree murder, and apologised for his actions, explaining that he had been in a rage. An example of an honest appeal in this sample, was a man who made an appeal for help in finding who had killed his daughter, who had been found raped and strangled in the driveway of her home. DNA evidence on her body

did not match any on the police database, but a year later, it was matched to DNA taken from a man arrested for a minor, unrelated offence. Dental records and fingerprints also linked the man to the murder, and he was subsequently convicted.

For studies 4, 5 and 6, a larger and different sample of stimulus materials was required, and so new appeals were collected. Internet searches were conducted with search engines looking for video content, using the following search terms: husband + plea, wife + plea, parent + plea, family + plea, husband + appeal, wife + appeal, parent + appeal, family + appeal, husband + missing, wife + missing, child + missing, missing + press conference, murder + press conference, family + press conference, appeal + press conference, plea + press conference, missing + plea, missing + appeal, murder + appeal, murder + plea.

All deceptive appeals that satisfied the criteria outlined above (ground truth established, not recent and high profile in the UK, and made recently after the event) were included in the sample; deceptive appeals were less numerous than honest appeals, and it was desirable to have as large a sample size as possible. The author became aware of the existence of several deceptive appeals that were not available online, unfortunately funds were unavailable to purchase the footage from news archives.

Thus 16 deceptive appeals were included in this sample, and honest appeals included were matched as closely as possible to the deceptive appeals in terms of: relative missing or publicly known to be dead (i.e. body found), and gender of the appealer. Thus the final selection included 16 deceptive appeals, of which 10 appeals involved missing relatives, and six appeals involved dead relatives. Sixteen honest appeals were included, of which 10 appeals involved missing relatives, and six



appeals involved dead relatives. A summary of the properties of the appeals is presented in Table 5.2 below

Table 5.2: *Properties of appeals used as stimulus materials in studies 4, 5 and 6*

	Relative missing	Relative dead	Gender	Length in seconds <i>M (SD)</i>
<b>Deceptive appeals</b> ( <i>n</i> = 16)	10	6	11 male 5 female	48.22 (34.57)
<b>Honest appeals</b> ( <i>n</i> = 16)	10	6	10 male 6 female	52.13 (36.08)

An example of a deceptive appeal in this sample was a woman who reported that she had been car-jacked and the perpetrator had driven away with her children in the car. The woman made a televised appeal for help to find her missing children, however inconsistencies in her account placed suspicion on her, and nine days after reporting her children missing, she confessed that she had rolled her car in to a lake with her children inside. The woman was able to tell police exactly where in the lake to find her car, in which the bodies of her drowned children were found. An example of an honest appeal in this sample, was a woman who appealed for help in finding the murderer of her husband, who had been stabbed 18 times. Two weeks later, a man handed himself in to police, admitting to the killing; the man he had murdered had been due to testify against him, and extensive eyewitness and forensic evidence linked him to the crime, as well as his confession.

Study 7 utilised the same 32 appeals as studies 4, 5 and 6, plus four ‘filler’ appeals. These ‘filler’ appeals, two honest and two deceptive, were selected at random from the available pool of appeals; three appeals used in studies 1, 2 and 3,

and one appeal not previously used, were included. A summary of the properties of the appeals is presented in Table 5.3 below.

Table 5.3: *Properties of appeals used as stimulus materials in Study 7*

	Relative missing	Relative dead	Gender	Length in seconds <i>M (SD)</i>
<b>Deceptive appeals (<i>n</i> = 18)</b>	10	8	12 male 6 female	48.92 (32.86)
<b>Honest appeals (<i>n</i> = 18)</b>	11	7	11 male 7 female	52.22 (34.14)

#### 5.1.1. *Ground truth*

As noted, to be acceptable for inclusion in the present studies, all appeals had to satisfy the criterion of adequate ground truth; i.e. the evidence was strong enough to support the conclusion that the appeals were actually truthful or deceptive. According to other published research in this area, a number of criteria can be used to determine whether ground truth has been established; those used by ten Brinke and Porter (2012), and Vrij and Mann (2001b), were used here.

Consequently, in the present research, appealers were classified as honest or deceptive only if there was overwhelming evidence indicating the extent of their involvement in the death or disappearance of their relative using these criteria. Such evidence included: forensic evidence (for example, soil traces, pollen traces, fibres linking the accused to the crime scene, blood spatter patterns, bite marks); presence of the victim's blood (for example, in the clothing, car, or home of the accused); other DNA linking the accused to the crime (for example, skin, hair, body fluids);

footage from security cameras, CCTV, police surveillance videos, and speed cameras (for example, CCTV of the accused dumping the body of the victim); knowledge of the location of the missing person's body; knowledge of unreleased or undiscovered details of the crime (for example, knowledge of the cause of death before the body was found by police); confessions which included intimate knowledge of the crime and were not recanted; admission at trial (for example, admitting guilt and apologizing to family members); post mortem evidence (for example, evidence that the victim could not have been alive at a time when the accused claimed that he or she was alive); medical evidence (for example, expert testimony that it was impossible for a person to be rendered unconscious for 10 hours by a drug administered in a way that the accused claimed); computer search history of the location of the crime or the body, prior to the crime occurring (for example, zoomed-in satellite images before the relative was reported missing, of the exact, remote location where the body was later found); eyewitness testimony (for example, a family member who survived an attack in which other family members were killed, or identification of the accused at the crime scene); possession of the murder weapon; possession of items from the crime scene; phone records; incriminating financial transactions; and an account of events, or an alibi, inconsistent with the evidence (for example, a man claiming to have spent the morning searching for his wife, when CCTV footage shows him taking a mattress to a dumpster, and till receipts show that he then purchased a new mattress. His wife's body was later found wrapped in a bloody part of a mattress). The large majority of cases involved multiple pieces of evidence as described above. Following the protocol established by ten Brinke and Porte (2012), Table 5.4 provides a summary of the types of evidence used to classify cases as deceptive or honest.

Table 5.4: *Frequency of case evidence used to establish ground truth*

Evidence type	Honest appeals ( <i>n</i> = 22)	Deceptive appeals ( <i>n</i> = 20)
Forensic evidence (soil, pollen, fibre, fingerprints, dental records, blood spatter patterns etc)	20	9
DNA evidence (body fluids, hair, skin etc.)	9	5
Victim's blood	4	5
CCTV/video/speed camera evidence	8	3
Knowledge of location of victim's body	3	3
Knowledge of unreleased/undiscovered details of crime	3	3
Confession	8	9
Admission at trial	7	9
Post mortem evidence	1	5
Medical evidence	0	1
Computer search history of location of crime/body	2	0
Eyewitness testimony	7	9
Possession of murder weapon (appealer/accomplice)	0	2
Possession of items from crime scene	2	1
Phone records	0	2
Wiretap evidence	0	1
Incriminating financial transactions	1	3
Account/alibi inconsistent with evidence	5	14

Note: totals exceed sample size as cases were classified based on several pieces of evidence

In total, 42 appeals were included in the final selection of cases, approximately half of which were classified as deceptive, and half as honest. As previously mentioned, not all appeals were used in all the studies.

In the group of appeals classified as deceptive, in all cases except one, the appellant has also been convicted in a criminal court of involvement in the death (or kidnapping, in one case) of their relative. The exception concerned a mother suspected of killing her missing child who led police to the child's body, buried in a remote location, whilst she was under surveillance. She has pleaded guilty to the manslaughter of her child, and to interfering with her remains. However, at the time of writing, her plea was not accepted by the court and she is awaiting trial charged with murder.

In the group of appeals classified as honest, another person has been convicted of the death of the appellants' relative, or the relative was found with no evidence of foul play, except in two cases. One of those two cases was that of a father who made an appeal for help when his daughter disappeared whilst in the care of her mother, from whom he was separated. CCTV footage from the day the girl was reported missing clearly shows a man, an acquaintance of the girl's mother, carrying the girl down a hotel corridor. The girl's body was found shortly afterwards. The man was charged with first degree murder, rape of a child and kidnapping, the mother was charged with murder, human trafficking, child abuse involving prostitution, and filing a false police report. From initial investigation to trial, the man has not disputed that he took the child, and information provided by the man's lawyer helped police find the child's body. At the time of writing, the trial continues. The father who made the appeal lived in a different state to his daughter, was never named as a person of interest or a suspect in the case, and the CCTV footage from the hotel clearly shows with whom his daughter was just before she died.

The second of those two cases (used only in Study 7) was that of a mother who appealed for the return of her teenage daughter, who had disappeared when

attempting to hitch-hike home from a rock concert. The daughter's body was found buried in a remote location three months later; she had been raped and violently murdered. DNA evidence linked the murder to a male suspect in a previous sexual assault in the same area some years previously. The victim in this case was able to get a good look at the man, enabling police to produce a sketch of the attacker, although his identity remains unknown.

## **5.2 Between Subjects Stimulus Material Design**

It can be noted that the present study adopted a between subjects stimulus materials design; i.e. it concerned comparing honest and deceptive individuals. In the majority of previous research investigating cues to deception in high stakes situations, a within subjects stimulus materials design has been used (see, for example, Davis et al., 2005; Mann & Vrij, 2006; Mann et al., 2002, 2006; Villar et al., 2011; Vrij & Mann, 2001a). In these studies, honest and deceptive sections of communications by individuals guilty of committing various crimes (commonly from police interviews) are compared. This design is limited in two main respects. First, there are many instances in real life where the evidence available is not sufficient to determine what parts of a person's statement are or are not credible. If the whole statement is a lie, or little is said, it is more or less impossible to use a within subjects method. Second, the within subjects method is particularly limited as it negates the possibility of examining the individuals holistically; i.e. if a person is lying in parts of his/her testimony, it does not follow that cues will be limited only to those parts of the testimony which are untruthful. For example, if lying involves cognitive load and the

experience of anxiety, there is no reason why these should be exhibited only when the person is stating something false; even when making some truthful statements, a person who is lying still has to monitor what he or she has said and intends to say, and anticipate what may be asked etc. It was, therefore, expected that the use of a between subjects stimulus materials design would facilitate the emergence of cues; indeed, two previous studies of high stakes deception using this design have found cues that are likely to be context-specific (Harpster et al., 2009; ten Brinke & Porter, 2012).

### **5.3 Limitations of the Stimulus Materials**

However, notwithstanding the advantages of using these kinds of video stimulus materials, there were a number of limitations attached to using them. Real world data is inherently ‘dirty’, and there are obvious problems relating to a lack of control over variables, associated with using such data. For example, the resolution of the videos was variable, so that some cues could not be investigated; blink rate has been found to both increase and decrease in high stakes deception (ten Brinke & Porter, 2012; Mann et al., 2002), however, the video images in the sample of appeals used in the present thesis, were often not sufficiently clear to investigate this cue. The scope of the image also varied in the appeals, so that in some cases the whole body of the appealer was shown, and in some cases just the head was shown; thus although illustrators have also been found to both increase and decrease in high stakes deception (Davis et al., 2005; Koper & Sahlman, 1991), again it was not possible to investigate this cue in the present thesis.

As noted all the appeals were collected from the United States, the United Kingdom, Canada and New Zealand, and the Western bias of the sample is acknowledged, as it was necessary to use English speakers to investigate verbal behaviours, and also observer accuracy. It is certainly possible that some cues may not generalise across cultures.

However, despite the limitations, because of their excellent ecological validity, it was considered that the advantages of such videos as stimulus materials hugely outweighed the drawbacks.



# **Chapter 6**

## **Study 1: A Qualitative Analysis of Cues to Deception Used by Participants Accurate at Detecting Deception in Appeals**

### **6.1 Introduction**

As discussed in the introductory chapters, factors considered to be important in the production of cues to deception (including cognitive load, affective responses, behavioural control and impression management, and arousal), may be exacerbated in high stakes situations. Consequently, cues to deception may be more prominent in high stakes situations. Furthermore, as previously noted, DePaulo and Morris (2004) have suggested that there may be cues to deception other than those already identified in laboratory research, and that as yet unidentified cues may emerge only in high stakes situations. Moreover, researchers have suggested that a variety of situational factors may affect the production and salience of cues related to deception (see, for example, DePaulo & Morris, 2004; Porter & ten Brinke, 2010; Sporer & Schwandt, 2006); thus factors underlying the production of cues to deception may be differentially influential in differing situations, so that even in high stakes situations,

all lies may not necessarily resemble each other. For example, in the specific context of making an appeal for help with a missing or murdered relative, cues related to affective responses may be more salient than cues related to cognitive load; an appeal could be considered to be a highly emotional event, but not necessarily highly complex cognitively, as appeals are not generally a detailed account of events.

The observation that there may be as yet unidentified, context-specific cues to deception that emerge only in high stakes situations, raises the question of developing a methodology to identify such cues. One possibility would be to investigate which cues are used by people accurate at detecting deception; if there are individuals who are more able than most to discriminate truth from deception, then it seems likely that the cues which they use to make their credibility judgements have some utility. Some limited attempts have been made to investigate the cues to deception used by individuals identified as unusually accurate at detecting deception, but this methodology is as yet underdeveloped. As noted in Chapter 3, there have been lively debates in the deception literature regarding the existence of individual variation in the ability to detect deception (see, for example, Bond & Uysal, 2007; O'Sullivan, 2007). However, studies in which real life, high stakes lies are used as stimulus materials consistently suggest that some individuals are better able than others, and much better able than chance, to detect deception; for example, researchers report individual accuracy rates as high as 100% (Vrij & Mann, 2001a), 90% (Mann et al., 2006), and 80% (Vrij & Mann, 2001b) in studies using high stakes lies as stimulus materials. It may be that individual differences in detecting deception become more salient when making credibility judgements about high stakes lies (Shaw et al, 2013), as it is in these situations in which cues to deception may become discernible.

As discussed in Chapter 3, the methodology of investigating cues used by accurate lie detectors is nascent. Some limited attempts have been made, but reported findings have been, on the whole, vague. Mann et al. (2006) asked police officers with high deception detection accuracy rates to select fragments from clips of suspect interviews that they found most useful, but did not report specifically which behaviours were considered important. O'Sullivan and Ekman (2004) have published a few preliminary and very general observations about how their 'wizards' are able to accurately detect deception, including a frequent use of nonverbal behaviour, inconsistencies/consistencies between verbal and nonverbal behaviour, and sensitivity to the nuances of language use, but no specific cues are reported. Bond's (2008) reports regarding the accurate deception detectors that he identified are equally general, including a high usage of nonverbal cues, and focusing on the face and the arm/torso area.

It appears that there is substantial scope for further research in this area, and that an inductive, bottom-up approach, identifying the cues used by accurate deception detectors, may be worth pursuing as a useful method for generating previously unidentified cues. Consequently, the first study described here was a qualitative investigation of the cues to deception used by individuals accurate at detecting deception in appeals, when making credibility judgements about those same appeals. The purpose of the study was to generate cues that may not have previously been identified and may be particular to the context of making appeals, to be tested in addition to more conventional cues, in future, quantitative studies. It was hoped that, notwithstanding previous research investigating cues used by accurate deception detectors, specific cues, rather than general observations, would be identified. As the nature of the study was not experimental but essentially inductive

and exploratory, no explicit hypotheses were proposed at the outset; however, because appeals for missing or murdered relatives are likely to be an emotionally charged event, it was expected that affective responses might be particularly salient. As perceptions of emotions in others are necessarily subjective, it was further expected that subjective impressions would be prominent in the cues suggested by the participants.

## **6.2 Method**

### **6.2.1. *Participants***

Two participants who had taken part in a previous study on cues to deception (Wright Whelan, 2009), and had both achieved accuracy rates above 87% in distinguishing between honest and deceptive appealers, were recruited for the study. Both participants were female, aged 30-40, and non-psychologists, and neither had received any training in lie detection. Participants for the previous study had been recruited using opportunity sampling, and the specificity of the sample is acknowledged. In the previous study from which the two participants were recruited, all participants had viewed 10 video appeals (five honest and five deceptive) and were asked to make credibility judgements about the appealers. The appeals used as stimulus materials are described in Chapter 5. The two most accurate participants correctly judged eight out of nine, and seven out of eight appeals, respectively (one of the participants was familiar with one case, the other was familiar with two of the cases). It is recognised that the sample size of stimulus materials was rather small,

however, as noted in Chapter 3, this is common in high stakes deception research (see, for example, Mann et al., 2006; Vrij & Mann, 2001a; Vrij & Mann, 2001b). It is also recognised that the accuracy of the participants was measured only once and was therefore not firmly established; however, the aim of the study was to generate possible cues to test in future studies; the validity of the cues was examined in later studies.

Data were also collected from the blog of one of the truth ‘wizards’ identified in O’Sullivan and Ekman’s (2004) programme, which has tested thousands of people and identified a handful who are able to reliably detect deception. The blog contains comments on cues to deception in real life cases as they occur (i.e. before the outcome is known), and includes comments on three of the appeals used in this study (Eyes for Lies, 2007, 2008a & 2008b). To re-iterate, the purpose of the study was not to establish the validity of the cues or the accuracy of the participants, but to generate cues to examine in future studies, and it was on this basis that data from the blog were included.

### ***6.2.2. Materials and Procedure***

Two participants had provided informed consent in the previous study, and ‘Eyes For Lies’ consented to the use of data from her blog. The two participants attended at a psychology laboratory at the University of Liverpool, and were interviewed individually. Participation lasted between one and a half, and two hours. The participants were informed that the aim of the study was to identify the cues that they had used to distinguish between honest and deceptive appealers in the previous

study, which had taken place several months before. The criteria for the selection of cases included in the study are discussed fully in Chapter 5. See Appendix 3 for case summaries of the appeals. The two participants were informed that they would watch the same video footage of people making appeals for missing or murdered relatives as they had in the previous study, and would be asked to make comments as they occurred to them, about anything that they had used to come to their decision of lying or truthful. Before viewing each appeal, the participant was told whether they had previously judged the appellant to be lying or truthful, but not whether their judgement had been correct or incorrect. Participants were allowed to 'pause' the appeals, and to watch each appeal as often as they felt necessary.

After viewing each appeal, a semi-structured interview schedule was used, to which the interviewer noted the participants' responses. The participants were first asked the general question 'What made you decide that this person was lying/telling the truth?' This question was asked to allow the participants to specify which cues were most relevant for them, and to allow for the possibility of uncovering previously unidentified cues. This general question was followed by probing questions in order to attempt to elicit specific cues that could be subsequently tested. This methodology of allowing the participants to comment freely, followed by probing questions, follows O'Sullivan and Ekman's research in to cues to deception used by participants accurate at detecting deception (2004).

If the participant did not mention emotion in answer to the first, general question, the interviewer then asked 'Can you describe the person's emotions?' This question was asked because in the specific context of people making appeals for missing or murdered relatives, it was expected that affective responses may be particularly salient. The participant was then asked 'Was there anything about the

person's facial expressions that helped you decide?' Facial expressions were of interest as the face is one of the primary modes of communicating emotion, and in a study by Porter and ten Brinke (2008), it was found that all participants leaked inconsistent emotional expressions when trying to produce deceptive expressions of emotion. This lends some support to Darwin's inhibition hypothesis (1872), that some aspects of genuine facial expression cannot be voluntarily simulated, but may be involuntarily expressed when a genuine emotion is felt; and also to Ekman's (2001) related proposal that during deception, emotional leakage in facial expressions is an involuntary and near-ubiquitous human behaviour. Moreover, a recent study investigating appeals for help with missing or murdered relatives, found that facial expressions related to disgust, distress, surprise and happiness discriminated between honest and deceptive appealers (ten Brinke & Porter, 2012), suggesting that this channel of communication is worthy of further investigation.

The participants were then asked 'Was there anything about the person's eyes that helped you decide?' The interest in whether the participants used information from the eyes as a cue to deception was based on previous findings that in the specific context of appeals, deceptive appealers were perceived to have less direct eye contact than honest appealers (Wright Whelan, 2009). Furthermore, gaze aversion has been associated with deception in a different high stake situation (Vrij & Mann, 2001a), and has been found in a meta-analysis to emerge as a cue to deception when moderated by motivation to succeed in the lie (DePaulo et al., 2003).

The participants were then asked 'Was there anything that the person said that helped you decide?' This probing question was intended to elicit any specific verbal cues that the participants used. The verbal content of speech has long been regarded by many researchers as being important in differentiating honesty and

deception (for example, DePaulo et al., 2003; Ekman, 2001; Steller & Kohnken, 1989; Shuy, 1998), as various factors associated with producing deceptive communication may result in cues such as speech disturbances, vagueness, and distancing strategies. Moreover, several studies investigating high stakes deception have found verbal cues to be useful, including evasive or equivocal language (Adams & Jarvis, 2006; Harpster et al., 2009; ten Brinke & Porter, 2012), speech dysfluency and errors (Davis et al., 2005; Vrij & Mann, 2001a), and word and/or phrase repetition (Davis et al., 2005; Harpster et al., 2009; Vrij & Mann, 2001a).

Next, participants were asked ‘Was there anything about the way the person spoke that helped you decide?’ This was intended to prompt any paraverbal cues used, as several studies that have examined high-stakes lies have reported that some paraverbal cues may be associated with deception. These include a higher vocal pitch (Hall, 1986; Wright Whelan, 2009), a lack of vocal modulation (Harpster et al., 2009), a less certain and direct voice (Koper & Sahlman, 1991; Wright Whelan, 2009), and increased pausing or hesitation (Mann et al., 2002; Vrij & Mann, 2001a; Wright Whelan, 2009).

Finally, the participants were asked ‘How did the person make you feel?’ This question was based on the finding that people feel significantly less comfortable when hearing a lie than when hearing a truth (Anderson, DePaulo, Ansfield, Tickle & Green, 1999), and on suggestions from leading researchers (Vrij, 2004; DePaulo & Morris, 2004) that implicit methods may be a promising area for future research in deception detection. It was also designed to investigate whether emotional contagion (Hatfield, Cacioppo & Rapson, 1992) was a factor in the participants’ judgements: it has been shown that viewing photographs of sad faces evokes a corresponding



affective response in the viewer (Wild, Erb & Bartels, 2001), and it may be that this effect is more likely to occur when viewing faces expressing genuine emotions.

The nine appeals used in the study were presented to the participants in two different randomised orders. There were five honest appeals and four deceptive appeals (as noted previously, one of the deceptive appeals used in the previous study was known to both the participants). The extracts used from the 'EyesForLies' blog in this study related to three of the deceptive appeals, and were all made when the appeals were first broadcast and before the outcome of the case was known.

### **6.3 Results and Discussion**

The responses of the participants and the data from the blog were analysed using the inductive thematic analysis procedure described by Hayes (2000) and Braun and Clarke (2006). The data were analysed as three separate elements; liars, truth-tellers and mistakes (both participants made one incorrect judgement on the veracity of the appealers). First, the responses were read carefully and repeatedly to identify units of text relevant to the research topic and search for meanings and patterns. Second, units relating to the same topic were grouped together in to categories and initial codes were generated. Units of text could be included in more than one category, and the entire content of the data set was coded. The data were reviewed to ensure that each category was sufficiently supported by the units.

The analysis identified five key themes for the deceptive appeals, and four key themes for the honest appeals, each with a number of associated categories,

resulting in 50 cues. There were two key themes for the mistakes. The themes and categories with numbers of units, and examples, are presented in Table 6.1.

It can be noted here that the conceptual distinction between cues to honesty and cues to deception is an important one. The premise underlying this distinction is that whilst the presence of a particular cue may probabilistically indicate honesty, its absence does not necessarily indicate dishonesty; similarly whilst the presence of a particular cue may probabilistically indicate deceit, its absence does not necessarily indicate honesty. Indeed, to infer one from the other would be to commit a logical fallacy of the form, ‘if A then B, therefore, if not A then not B’. This is like saying, if the Sun is shining then it is daytime, therefore, if the Sun is not shining, then it is night time; or if you join Alcoholics Anonymous you are likely to be a drinker, therefore, if you are not a member you must be teetotal.

### **6.3.1. *Deceptive Appealers***

Analysis of the responses to the deceptive appealers resulted in five key themes: fake emotion; lack of emotion; distancing behaviours; implausibility; and negative personal reaction.

Table 6.1: *Qualitative analysis; themes and their composite categories (number of units relating to category in brackets), with examples*

## **Deceptive appealers**

### **Fake emotion**

#### **Fake emotion (11)**

*E.g. 'you can tell she's not really upset, she doesn't get emotional although she keeps trying to look as if she's about to, and she's not crying'*

#### **Fake facial expression (4)**

*E.g. 'She's trying her best to look upset, but it's fake'*

#### **Putting on a performance (6)**

*E.g. 'Looking to the audience, like he was performing.'*

#### **Pretending to cry (6)**

*E.g. 'If he is so upset, why aren't there tears?'*

#### **High vocal pitch (15)**

*E.g. 'that constant high pitch is fake'*

#### **No emotional variation (4)**

*E.g. 'his emotions are at a constant level – it's put on'*

### **Distancing**

#### **Impersonal (7)**

*E.g. 'she would want to speak from the heart to her daughter, but she doesn't get personal in any way'*

#### **Focus on others (9)**

*E.g. 'He was going on about a lot of other people's losses – if it was my wife, I wouldn't give a f\*\*\* about anybody else's loss.'*

#### **First person pronouns (2)**

*E.g. 'She says "It's hard not to know where your kid is." To refer to her daughter as "your kid" is strange and removed'*

#### **Equivocation (3)**

*E.g. 'You expect a definitive answer, but it's half-baked.'*

#### **Irrelevant/generic statements (2)**

*E.g. 'He makes the most generic statements'*

#### **Not focused on relative (8)**

*E.g. 'she hasn't yet mentioned the loss of her children'*

#### **Lists (7)**

*E.g. 'All the lists of people are unnecessary'*

#### **Gaze aversion (8)**

*E.g. 'he doesn't make eye contact with anyone'*

### **Lack of emotion**

#### **Lack of emotion (6)**

*E.g. 'total lack of emotion talking about her children'*

#### **Painless recall (7)**

*E.g. 'there's a total lack of emotion talking about her children and what they suffered'*

#### **Calm/calculated (9)**

*E.g. 'her list of 'Call the doctors, the blood type, call the cops' is too calculated and calm – too much thought has gone in to it'*

#### **Lack of emotion in voice (4)**

*E.g. 'Her voice is calm'*

#### **Emotionally cold (7)**

*E.g. 'there's something unloving about her'*

#### **Lack of facial expression (4)**

*E.g. 'there was no expression of sadness at all'*

#### **Brutal language/detail (3)**

*E.g. 'She's very calm talking about blood coming out of her daughter's mouth'*

### **Implausibility**

#### **Nonsensical (9)**

*E.g. 'nothing about his plea is logical'*

#### **Making assumptions (2)**

*E.g. 'she acts like her daughter won't be able to see the TV.'*

#### **Lack of hope of finding missing relative alive (7)**

*E.g. 'Why would he give up hope for her so soon? Why would he assume the worst without any evidence to suggest otherwise unless he knows her fate?'*

#### **Indication of dispute (2)**

*E.g. 'It's obvious from what he tells us that they had a fight'*

#### **Defensive (3)**

*E.g. 'she's defensive'*

#### **Fear/nervousness (9)**

*E.g. 'He's nervous – that wavering voice and twitchy hand at his throat'*

### **Negative personal reaction**

#### **Dislike for the appellant (6)**

*E.g. 'I didn't like her at all'*

#### **Creepy (2)**

*E.g. 'She gives me the creeps.'*

#### **Behaving unnaturally (3)**

*E.g. 'She's unnatural and false'*

#### **No sympathy for the appellant (6)**

*E.g. 'I didn't feel anything for her'*

---

## **Honest appealers**

### **Genuine sadness**

#### **Genuine sadness (8)**

*E.g. 'it was all really genuine and heartfelt, she's not putting anything on'*

#### **Facial expressions (4)**

*E.g. 'she looks crumpled and destroyed'*

#### **Sad eyes (5)**

*E.g. 'She's got sad eyes'*

#### **Desperation/urgency (4)**

*E.g. 'she had an overall sense of desperation'*

### **Genuine/heartfelt**

#### **Genuine/heartfelt (7)**

*E.g. 'it came across as real and authentic'*

#### **No prepared speech (5)**

*E.g. 'she was speaking from the heart, not a prepared speech'*

#### **Positive emotion expressed towards relative (6)**

*E.g. 'He's proud of his daughter, speaking about her and not about himself'*

#### **Smile appropriate to speech content (2)**

*E.g. 'He has a small smile when he's talking about her and describing the way she was'*

#### **Genuine behaviours (3)**

*E.g. 'looking up to the ceiling when he was remembering her, really visualising her, that was genuine'*

#### **Credible eye contact (6)**

*E.g. 'She was looking around at everyone and down every now and then – indicates that she's legitimate'*

### **Containing emotion**

#### **Containing emotion (15)**

*E.g. 'She's containing her emotions for the sake of her daughter'*

#### **Calmness appropriate to speech content (9)**

*E.g. 'she's being calm to get her daughter back'*

#### **Trying to be sensible (8)**

*E.g. 'She's thinking about it in as sensible a way as you can'*

#### **Hopeful of finding missing relative (2)**

*E.g. 'she has hope that her sister's alive'*

#### **Avoiding brutal detail (2)**

*E.g. 'he doesn't talk about the body or anything'*

### **Sympathetic personal reaction**

#### **Feel the appealers' pain (3)**

*E.g. 'I could feel that she was upset'*

#### **Feel sorry for the appellant (7)**

*E.g. 'I felt sorry for her'*

#### **Normal (3)**

*E.g. 'He was just a normal bloke'*

#### **Behaving as expected (4)**

*E.g. 'it's what you'd expect'*

## **Mistakes**

### **Emotion not genuine**

#### **Performance (2)**

*E.g. 'she was making a speech'*

#### **Gaze (3)**

*E.g. 'looking around at everybody like she was giving a speech'*

#### **Inconsistent emotion (2)**

*E.g. 'she starts off looking upset but then is composed'*

### **Lack of Emotion**

#### **Lack of emotion (4)**

*E.g. 'he's very together for someone mourning a daughter'*

#### **Prepared speech (4)**

*E.g. 'it sounds like he's reading a letter'*

#### **Lack of grief (6)**

*E.g. 'it's too balanced for the immediate aftermath'*

---

#### 6.3.1.1. *Fake emotion*

A major source of cues to deception for the participants was emotion recognised as not being authentic. There were general comments about fake emotion (11) relating to two of the four lying appealers, for example ‘he does show emotion, but it’s totally false’, and ‘She doesn’t seem upset but is trying to look it – it’s fake.’ Related to this was the perception of the appellant as giving a performance (6), for example ‘It’s like he was performing, giving a speech.’ As well as these general perceptions, there were specific behaviours that the participants used to judge fake emotions. Pretending to cry, specifically, attempting to appear very distressed but producing no tears (6), was cited as a cue to deception, for example ‘That gesture of covering her eyes as if unable to go on, but there’s no emotion with it, she’s not really crying.’

Vocal pitch (15) was a specific indication of fake emotion, particularly for one appellant, for example ‘he pitches his voice up so high in what sounds like a fake attempt to show distress.’ The use of this cue is consistent with Hall’s (1986) research in which he measured the pitch of lies and truths told during criminal investigations, and found that in this high-stakes situation, liars had a higher vocal pitch than truth-tellers. In the context of deception, high vocal pitch is associated by Vrij (2004) with arousal, by DePaulo et al (2003) with tension and anxiety, and Ekman (2001) suggested that the fear of apprehension associated with high-stakes lying would result in higher vocal pitch. It is interesting that in the present study the participants judged a high vocal pitch and lack of modulation not as an indication of an affective state (such as anxiety), but as an indication of a simulated affective state.

Fake facial expression (4), including a lack of variation in expression, was used as an indicator of fake emotion, for example ‘Her eyebrows were always raised, a frozen surprise expression. It’s interesting to see expressions people will put on

when they want to fake sadness.’ This is in line with Porter and ten Brinke’s (2008) finding that people find it difficult to simulate convincing displays of negative emotions, including sadness, and Ekman’s (2001) suggestion that expressions of sadness are particularly difficult to fabricate, as they involve raising and drawing together the inner corners of the eyebrows, something which most people cannot do voluntarily. In particular, this relates to a recent study by ten Brinke and Porter (2012) which analysed the facial expressions of people making honest and deceptive appeals, frame by frame, using a facial coding system. The researchers found that honest appealers were more likely to express upper face sadness, whereas deceptive appealers were more likely to produce expressions of surprise, characterised by raised eyebrows, which the researchers speculate was a result of failed attempts by the deceptive appealers to simulate sadness.

A lack of variation in emotion (4) was also used as an indicator of fake emotion: ‘There’s no variation. His face is stuck in that expression with that monotone high pitch.’ It may be that a lack of modulation in the voice, facial expression and perceived emotional expression is particularly salient as pointing to fake emotion.

#### 6.3.1.2. *Lack of emotion*

The second key theme and major source of cues to deception for the participants was a lack of emotion, relating particularly to two of the deceptive appealers (fake emotion was related almost entirely to the other two deceptive appealers). There were general comments about a lack of emotion (6), for example ‘There was no emotion’, and ‘He didn’t seem to be emotional.’ As well as these general perceptions, there were also more specific behaviours that the participants used to

judge a lack of emotion. Recalling events related to the loss of their relative without pain (7) was a cue used by both participants, for example ‘to witness something like that happening to your children, you’d be distraught talking about it, but she is calm.’ This observation is in line with research showing that recall of emotional events involves re-experiencing the affective state (for example, Damasio et al, 2000; Dolcos, LaBar & Cabeza, 2005).

The perception of the appellant as being calm and calculated (9) was an important cue in the theme of lack of emotion, for example ‘Everything about her is as if it’s thought out’, and ‘To be as calm as that – if she died naturally, but to be horrifically killed – he didn’t seem quite genuine.’ A lack of emotion in the voice (4) contributed to the perception of a lack of emotion, for example ‘A very calm and even tone, his voice doesn’t break. It’s as if he’s having a conversation’, and ‘her voice was totally emotionless.’ This is in contrast to the deceptive appellants who were judged more as having fake emotions, where a high vocal pitch was reported as a cue to deception.

A general perception of coldness (7) was a category that emerged as contributing to a lack of emotion, for example ‘She was so calm and cold, there was a total lack of empathy for her children’, and a lack of facial expression (4) ‘There weren’t any facial expressions. She was blank, unemotional’. Related to coldness was the use of brutal language or detail about the relative (3), for example, ‘I wouldn’t be able to verbalise it, such a horrific injury to my child’, and ‘he uses the word "murder". It's a strong word’. In Criteria-Based Content Analysis (Steller & Kohnken, 1989), the quantity of detail and the use of unusual detail are regarded as indicative of truthfulness, but in the specific situation of making an appeal for a relative, the use of vivid, brutal detail in relation to a relative was flagged as a cue to

deception; appealers who were able to use language like this when talking about their relative were regarded as showing a lack of appropriate emotion in this context.

#### 6.3.1.3. *Distancing*

The theme of the appealer distancing him/herself from the missing relative was relevant to all four deceptive appealers, whether they were regarded as displaying fake emotion or as showing a lack of emotion. An impersonal appeal (7) indicated to the participants that the appealer was not genuine, for example, 'Usually in appeals, people talk personally about the loss of someone they love, he doesn't.' This category is related to focusing on others (9), for example 'When he says "It's not just my loss, it's friends losses, it's student's losses, it's neighbours losses", the statement lacks personal connection. It's almost as if he's distancing himself.' Two possible specific functions of focusing on others emerged: using other people as a vehicle for appropriate emotions, for example, 'Putting other people before her on the list and attaching emotions to others, like "your friend is upset"', and as a respite, 'Dwelling on everyone else and thanking others is like a break from putting on fake emotion.'

The use of pronouns (2) was a cue to deception for the participants in one of the deceptive appeals, in which the appealer was noted as not referring to herself in the first person:

She says in her plea to her daughter, "Your Aunt is here. Your mom is here."

She doesn't say "It's me, mom." She doesn't refer to herself as I. She doesn't say "I" miss you, "I" love you. Instead, she says "we".'

There is a body of research (for example Newman, Pennebaker, Berry & Richards, 2003; Dilmon, 2009) supporting the participants' suspicion that a lack of self-



references may be related to deception. One deceptive appealer was noted as using equivocation (3), which was considered suspicious,

he says "...everyday it gets *a little bit* harder *in some ways*...But it's *kinda* days like today or moments like this..." These statements show that he's having some confusing and conflicting feelings.'

The use of irrelevant or generic statements (2) was another linguistic cue suggestive of distancing, for example, 'she's unbelievable, talking about her "entrapment by society."' Both the use of equivocal language, and generalising by using generic statements, are in line with research suggesting that liars display more non-immediacy in their language, making their speech more nonspecific and tentative (Zhou, Burgoon, Twitchell, Qin & Nunamaker, 2004), and with De Paulo et al's (2003) meta-analytic finding that deception is associated with linguistic constructions that distance the liar from the subject of their speech. As noted earlier, the use of equivocal language has been associated with deception in previous research on high stakes lies.

A lack of focus on the victim (8) contributed to the distancing behaviour of deceptive appealers, for example, 'Despite her three kids being shot, she's talking about herself', and 'She's thanking businesses instead of being focused on her daughter.' This may in part reflect a lack of urgency for resolution of the case on the part of the deceptive appealer. The emergence of the use of lists (7) as a cue to deception in the specific context of making appeals was unexpected, and was interpreted by the participants in two ways. Firstly, for one of the deceptive appealers, both participants saw listing as an indication of premeditation, 'She listed things - the doctor, blood group and cops – it's very logical and calculated. Lists are about being in control'. Secondly, listing was seen as a failed attempt to appear

genuine, 'She lists everybody who is missing her daughter, trying to think of things to say to sound normal', which in fact betrayed a lack of urgency, 'she's too busy listing rather than saying anything relevant to finding her daughter.'

For two of the deceptive appealers, gaze aversion (8) was a distancing category that was used as a cue to deception, for example 'There's no eyeline, he looks down the whole time.' 'Eyes For Lies' was more specific in her use of gaze aversion as a cue to deception,

She pleads to her daughter to make contact, yet she doesn't look at the camera. She looks down or away. Then later, after she finishes talking to her daughter, she looks back up normally as she talks. It is this distinction in her behaviour that is a big red flag.

Here it is the specific timing of the gaze aversion that is relevant; while the appealer is directly addressing her daughter, she is unable to look at the camera. This may be a cognitively driven behaviour, as the appealer knows that the person she is addressing cannot see her (she had already killed her daughter), and it may also have affective components as downward gaze has been related to shame (Keltner & Buswell, 1996), and to guilt (Ekman, 2001). As noted in Chapter 2, gaze aversion as a cue to deception has been dismissed as a popular myth (for example, Vrij, 2004), and yet has been found to be significantly related to deception when motivation to succeed in the lie was a moderating variable (De Paulo et al, 2003), and in one study of high stakes deception (Vrij & Mann, 2001a). It may be that the specific context of the lie is a critical element in whether or not gaze aversion is related to deception, and this behaviour clearly needs further investigation.

#### 6.3.1.4. *Implausibility*

The theme of implausibility related almost entirely to verbal behaviours. The perception of an appeal as being nonsensical (9) was used as a cue to deception, both in general terms, for example, ‘nothing about his plea makes sense’, and also more specifically; in the one case amongst the deceptive appealers where the relative admitted to being present at the event (she claimed that a stranger had shot her and her children), the point of view she used to describe the events was noted as not making sense ‘She’s describing it from the point of view of the person doing it.’ Related to the perception of appealers not making sense, was making assumptions (2), ‘She also acts like her daughter won’t be able to see the TV. If you don’t know where your daughter is, how can you make this assumption?’ Two of the deceptive appealers had a relative who was reportedly missing, not dead (i.e. no body had been found), and their lack of hope of finding the relative alive (7) was used by the participants as a cue to deception, for example ‘She said “Nothing can be done to make things better.” Why can’t things be done to make things better? Why wouldn’t finding her daughter make things better?’ Relatives of missing people are clearly expected to hang on to every hope until a body is found. These findings are consistent with De Paulo et al’s (2003) meta-analytic finding that lies make less sense than truths, being less plausible, less logical and sensible, and more likely to be internally discrepant or ambivalent.

An indication of dispute (2) alerted the participants to the possible guilt of the appellant in one case, ‘He says “She can have my car. She can have everything.” They’ve obviously had a row’, and a defensive attitude (3) was noted by the participants in another, ‘she’s making excuses and defending herself.’ The final category in the theme of implausibility was fear/nervousness (9), emotions that the

participants did not consider plausible in a genuine appeal, for example, 'He looks frightened, really frightened. Why would he be afraid?' Fear and nervousness were revealed facially, 'I see fear in his eyes and in his face', vocally, 'He's nervous. That wavering voice', and through gesture, 'He's twitchy with his hand at his throat', and 'she's plucking at her throat.' People's use of self-touching as a cue to deception has some general support from research in to nonverbal communication. For instance, Ekman and Friesen's (1969) affective explanation for self-touching postulates that self-touching releases emotional arousal and is used to self-comfort, but it is unintended, and thus may betray aroused affective states including anxiety, guilt and hostility. Harper, Wiens and Matarazzo (1978) also found that self-touching correlated with levels of being upset, and was particularly associated with hostility and suspicion. Also of relevance here are Freedman's (1972) psycholinguistic perspective, in which self-touching is suggested to occur when there is a conflict between what has been verbalised and what remains unsaid, and a more recent finding of a correlation between shame and self-touching (Gieler et al, 2008).

#### *6.3.1.5. Negative personal reaction*

Interestingly, both participants had strong negative personal reactions to all the deceptive appealers. The deceptive appealers were disliked (6), for example 'I didn't like her and I didn't like listening to her appeal', and 'I didn't like him at all', and were judged to be creepy (2), for example, 'He's creepy, there's something about him but I'm not sure what.' Related to this was the perception of lying appealers as unnatural (3), for example, 'Everything about her is as if it's thought out and unnatural', and 'He's weird.' This is in line with De Paulo et al's (2003) meta-analytic finding that liars are judged as less pleasant than truth-tellers.

The lying appealers also failed to elicit sympathy from the participants (6), for example, 'It leaves me cold. Usually hearing appeals from parents, my heart goes out to them, but it didn't here. I didn't feel anything for her.' This lack of sympathy occurred even when reasons for honesty were explicitly stated, for example, 'I didn't feel sorry for him, although the speech content was believable', and 'I didn't get any emotion from her. The physical gestures are there, eyes downcast and voice breaking, but it left me cold.' These findings are interesting in relation to the theory of emotional contagion mentioned earlier (Hatfield et al., 1992), in which a corresponding affective response to emotion displayed by another person occurs in the viewer. In the present study, emotional contagion clearly did not occur if the emotion expressed was not genuine; instead a lack of emotional contagion seemed to become a cue to deception in the context of appeals.

### **6.3.2. *Honest Appealers***

Analysis of the responses to the honest appealers resulted in four key themes: genuine sadness, containing emotion, genuine/heartfelt, and sympathetic personal reaction.

#### **6.3.2.1. *Genuine sadness***

Unsurprisingly, honest appealers were described in general terms as displaying genuine sadness (8), for example, 'It looks like he's got the weight of the world on his shoulders.' As well as general impressions, specific behaviours were used to judge the authenticity of the sadness of the honest appealers. Facial expression (4)

was used as a cue in two of the appeals, for example, ‘She’s not putting anything on, the look on her face is troubled’, and in one case before the footage started,

Just looking at her face, her face looks really, really sad and desperate. She looks destroyed. Is her daughter already dead? She looks like she’s got no hope, she’s destroyed.

In this case, the participant was correct in thinking that the body of the appellant’s relative had already been found. This is in strong contrast to the lying appellants who were unable to convince the participants that their facial expressions of sadness were genuine. Sad eyes (5) were an important specific component of a genuinely sad expression, for example, ‘mostly the expression in her eyes, genuine sadness’ and ‘Her eyes looked really sad. I don’t know how to describe it but her eyes were genuinely sad.’ Although the participants were unable to describe exactly what made the eyes look sad, they were commenting on appellants whose eyebrows were raised and drawn together at the inner corners; as noted earlier, this is an expression that is difficult to achieve voluntarily (Ekman, 2001), and which discriminated between honest and deceptive appellants in an analysis using a facial coding system (ten Brinke & Porter, 2012)

The final category that contributed to the theme of genuine sadness was a sense of desperation or urgency (4), whether or not a body had been found, for example, ‘She’s begging for the return of her daughter’, and ‘the way she said “you *must* (emphasis) call the police”, the “must” was genuine and credible, there was urgency, an emotion you’d expect in an appeal.’ This is in contrast to some of the lying appellants, whose failure to be appropriately focused on their relative was used as a cue to deception.

#### 6.3.2.2. *Containing emotion*

The theme of containing emotion contained five categories as well as general comments (13), for example, ‘She looked sad but not overly emotional, she’s trying to keep it together’, and ‘She looks like she’s been very upset but isn’t putting it on – there are no false tears.’ Interestingly, none of the genuine appealers cried, compared to two of the lying appealers who attempted to simulate crying. An appropriate calmness (9) was noted in three of the honest appealers, and it is important to note that each time it was mentioned, the participants specifically noted that the calmness was justified by the speech content, for example, ‘she’s keeping calm because she has to make an appeal to get her daughter back’, and ‘At the beginning he’s describing the way she was, he’s got nice memories, which explains why he’s not too upset.’ This contrasts with the deceptive appealers, where the category of calm/calculated contributed to the theme of lack of emotion, and was related to a more pervasive calmness or an inappropriate calmness.

Trying to be sensible (8) was a category that contributed to the theme of containing emotion, in relation to the two honest appeals in which the relative was missing (not known to be dead), for example, ‘she’s being rational at the beginning.’ This was related to being hopeful, for example, ‘there’s a sense of hope at the moment. She says “bring back in one piece” which indicates that she has hope that her sister’s alive.’ This contrasts with the deceptive appealers whose relatives were missing, both of whom displayed a lack of hope that was used as a cue to deception by the participants. In an honest appeal in which the body had been found, clearly holding out hope would not be possible, but it was noted that the appellant did not use brutal language (2) in relation to his relative: ‘He doesn’t go in to further details,

gruesome ones.’ This is in clear contrast to the cue of brutal detail/language used by the participants for some of the deceptive appealers.

#### 6.3.2.3. *Genuine/heartfelt*

All the honest appealers were described as genuine in general terms (7) by at least one of the participants, if not both, for example ‘It just all comes across as really genuine’, and ‘She’s talking about her grief.’ This is in contrast to the overall perception of giving a performance, used by the participants as a cue when looking at some of the deceptive appeals. Specific cues were identified as pointing to authenticity, including an unprepared speech (5), for example ‘she wasn’t reading, so it was really from the heart.’

The category of describing the victim positively (6) is in some ways related to, although distinct from, honest appealers not using brutal language or detail in relation to their relative, for example, ‘it came across that he loved her because of how he spoke about her, he was proud of her’, and ‘She’s talking about her lovely daughter.’ Dependent on this category is the category of appropriate smile (2), as smiling when describing the victim positively was considered a genuine behaviour, and was directly compared by one participant to a deceptive appealer who smiled when describing an alleged attack on her child, ‘He has a small smile when he’s talking about her but it’s a fond smile, not sinister, not like that other woman.’ It is a crucial point that the context of any behaviour must be considered; it may seem counter-intuitive that smiling during an appeal could indicate innocence, and indeed in some cases it would rightly arouse suspicion, and this suggests that simply counting and summing instances of behaviours may be misleading; they must be considered in relation to the specific moment in which they occur.



Examples of genuine behaviours (3) contributed to the overall impression of credibility of the honest appealers, for example, ‘That intake of breath, he’s getting himself together to continue talking.’ Related to this was the category of credible eye contact (6), which all the honest appealers were judged to have by at least one of the participants. It is important to note that although direct eye contact was judged a sign of honesty, for example, ‘He maintains eye contact’, it was not the only type of eye contact regarded as credible, for example, ‘She was looking out, then down momentarily, what you’d expect.’ Even more specifically, it was noted of the only honest appellant to directly address her relative:

She had a very direct eyeline to her daughter when she was speaking to her – if she’d looked down when she addressed her daughter she wouldn’t have seemed genuine.’

This contrasts directly to the deceptive appellant who addressed her daughter and was noted to direct her gaze downwards when doing so.

#### 6.3.2.4. *Sympathetic personal reaction*

The final theme that was identified in the participants’ comments about honest appealers was a sympathetic personal reaction. The participants felt the pain (3) of the honest appealers, for example ‘I could feel that she was really sad and upset’, and ‘I felt her pain’, and felt sorry (7) for all the honest appealers, for example, ‘it tugs at your heartstrings’ and ‘I felt so sorry for her. Poor, poor woman.’ It is notable that all the honest appealers elicited sympathy from at least one participant, and that not a single deceptive appellant elicited sympathy or feelings of sorrow. Of relevance here are not only theories of emotional contagion (Hatfield et al., 1992), but also recent research on mirror neurons, which suggest that we have neural mechanisms which

enable a form of direct experiential understanding of what we observe in others, which permits us to comprehend the emotions of others by experiencing them in our own brains (Rizzolatti, Fogassi & Gallese, 2006; Singer et al., 2004). Leading researchers in the field have commented that “When people use the expression ‘I feel your pain,’ they may not realize how literally it could be true” (Rizzolatti et al., 2006).

Two final categories identified in the theme of sympathetic personal reaction, were the appealers being regarded as normal (3), for example, ‘He looked like a normal bloke’ and ‘She didn’t give me the creeps’, and behaving as you would expect (4), for example, ‘The emotions you’d expect in an appeal’, and ‘It was what you’d expect when seeing an appeal.’ These categories are in direct contrast to the categories identified in the deceptive appealers of creepy and unnatural, and this may relate to norm violation: Levine et al. (2000) found that aberrant behaviour was judged to be more deceptive than normative behaviour whether or not it was deceptive, but it may also be that deceptive behaviour violates certain norms and appears aberrant (creepy and unnatural).

### **6.3.3. *Mistakes***

Both participants made one mistake and judged an honest appealers to be lying. Analysis of the responses made in relation to the two mistaken judgements resulted in two key themes; a lack of emotion and emotion not genuine.

#### 6.3.3.1. *Emotion not genuine*

The theme of emotion not being genuine contained three categories. One of the honest appealers was judged by one of the participants as putting on a performance (2), ‘She looked like she was putting on a show.’ Related to this was the use of gaze (3), which one participant thought was not credible, ‘She was looking around, making a speech.’ The last category in this theme was inconsistent emotion (2), a cue one of the participants used to wrongly judge one of the honest appealers, ‘She regained her composure quickly, after looking upset at first.’

#### 6.3.3.2. *Lack of emotion*

The theme of lack of emotion contained two categories as well as general comments (4), for example, ‘there’s no horror or shock, which I’d expect if a daughter had been murdered’, and ‘I didn’t feel that as a mother she was upset enough.’ Specifically, in the case of one of the honest appealers, the use of a prepared speech (4) misled one of the participants, ‘it’s like he’s reading an essay rather than talking from the heart’, although she did acknowledge the difficulty this presented, ‘this is more difficult to judge with prepared speeches’. Both appealers who were wrongly judged to be deceptive were regarded as displaying a lack of grief (6), for example, ‘She seemed a bit hard’, and ‘He’s already talking about getting on with their lives.’

### 6.4. General Discussion

The purpose of the present study was to generate potential context-specific cues to deception; a small number of the cues suggested by the participants have been

investigated elsewhere (for example, high vocal pitch, the use of pronouns, equivocation, and gaze aversion); some of the cues suggested were in line with, or could be predicted by, previous research; and a large number of the cues had not previously been identified. As expected, emotion emerged as a key element by which appealers were correctly judged as deceptive or honest, according to the participants in the present study. Deceptive appealers tended to be judged as either expressing fake emotion, or as displaying a lack of emotion, whereas honest appealers were judged as showing genuine sadness whilst also attempting to contain their emotions. The use of simulated emotion by deceptive appealers is almost certainly intended by them to mimic a genuine appeal, yet none of the respondents were convinced by any of the displays of fake emotion. Whereas some of the comments regarding fake emotion were general impressions, specific cues were also mentioned; a lack of tears when apparently crying, an unmodulated high vocal pitch, a lack of variation in emotion, and fake facial expression. In the honest appealers, specific signs of genuine sadness were heavily focussed on facial expression, particularly around the eyes, indicating that genuine expressions of sadness may be important in differentiating between liars and truth-tellers in the context of appeals.

The lack of emotion displayed by two of the deceptive appealers may be due to a genuine indifference to their relatives (perhaps indicative of a personality disorder), or may be explained in terms of an attempt at behavioural control resulting in affective rigidity (Vrij, 2004). Again, many of the comments relating to this theme were general in nature, although specific cues that could be subject to further investigation were suggested; the use of brutal language or detail, a lack of emotion in the voice, and a neutral facial expression. The participants were mostly successful in distinguishing between a lack of emotion in the deceptive appealers, and emotion

being contained by the honest appealers, although in both the cases in which the participants wrongly judged an appealers, they cited a lack of emotion (specifically grief) as a cue. This suggests that there may be an overlap in the behaviours that is a source of potential error, even for accurate deception detectors.

The use of distancing behaviours as a cue to deception is in line with Harpster et al's (2009) findings in their analysis of real life 911 homicide calls, in which verbal behaviours associated with evasion and distancing predicted guilt, and those associated with urgency and immediacy predicted innocence. Behaviours associated with distancing formed a theme that may be particularly useful in distinguishing deceptive appealers from honest appealers for two reasons. First, distancing behaviours were used by the deceptive appealers perceived to be faking their emotions, and by the deceptive appealers perceived to be lacking in emotion. Second, they are behaviours that may be more easily measured than many of the more impressionistic cues used by the participants. Statements that are impersonal, focus on others, are irrelevant or use generic phrases, are not focused on the victim, contain lists, have a lack of first person pronouns, or use equivocation, can all be quantified, as can gaze aversion. The theme of implausibility also contained categories of cues which lend themselves more readily to analysis; nonsensical, making assumptions, expressing a lack of hope, and indication of a dispute. The advantage of appealers who had not prepared a speech, or had not been assisted by the police in preparing a speech, is particularly evident here.

An unexpected theme for cues used by the participants to judge deception in appeals, was a negative personal reaction. Both participants either disliked, or had no sympathy for, all the deceptive appealers. This may have been an emotional response by the participants to their perception that the appealers were either fake or lacking in

emotion, a result of the belief that the appellant had killed their relative, or some other variable. More research is needed to investigate the cause or causes of this negative personal reaction; for instance, would people who believed the lying appellants nevertheless feel negatively about them? Substantially more data were produced by the participants on cues to deception (181 units), than on cues to honesty (103 units), and it may be that deceptive appellants violate norms of behaviour that are unnoticed when they are conformed to, and that this norm violation elicits a negative personal reaction.

Key contrasts that emerged from the data between deceptive and honest appellants were: in the cases where a body had yet to be found, the lack of hope expressed by the deceptive appellants contrasted with the honest appellants' hopefulness; the lack of focus on the victim in the deceptive appeals contrasted with the urgency in the genuine appeals; the impersonal and generic statements of the lying appellants contrasted with the honest appellants describing their relative positively; the lack of facial expression or fake facial expression in the lying appellants contrasted with the recognition of genuinely sad facial expressions in honest appellants; and personal reactions of a lack of sympathy towards the lying appellants contrasted with an empathic reaction to the honest appellants.

A crucial aspect of cues to deception in the context of appeals that emerged from the participants' responses was the relationship between behaviours. Much research has tended to assess cues to deception separately by counting occurrences and then summing them together without considering the dynamic interaction between cues, and yet the participants in the present study used a holistic approach that considered behaviours simultaneously and assessed cues as embedded in a context. For example, a lack of tears was used as a cue to deception but only in the

context of an appealer attempting to appear very distraught; none of the honest appealers cried, but their lack of tears was not seen as indicative of deception as it was not accompanied by feigned distress. Calmness was used as a cue to both deception and to honesty, depending on the context; calmness in an appeal in which the body had been found was considered suspicious, but in cases where the relative was only missing calmness as an aspect of containing emotion was considered appropriate behaviour.

Two important examples of considering verbal and nonverbal behaviour interactively instead of in parallel, were the participants' use of smiling and gaze aversion as cues. Smiling was considered to be indicative of genuine behaviour if it occurred when talking about a relative in positive terms, but not if it occurred during what should have been a painful memory. Gaze aversion was used as a cue independent of context in one case where it was particularly marked, but otherwise it was noted as a cue specifically in the context of the appealers' accompanying verbal behaviour, particularly when addressing their relative. Although Porter and ten Brinke (2010) have advocated this type of holistic approach, with concurrent attention to multiple channels of behaviour, they have also acknowledged that some manifestations of deception are idiosyncratic and difficult to address in quantitative research, so that sharp observation and psychological insight may be required as well as evidence from empirical research on high stakes lies, to accurately detect deception.

The present data could be construed as endorsing the view that different cues are likely to be relevant for different types of lies. As De Paulo and Morris (2004) suggest, the use of ecologically valid stimulus materials may produce previously unidentified cues that warrant further investigation, but at least some of these are

likely to be context specific. Lists and focusing on others may be useful cues to deception in deceptive appeals, but are unlikely to be useful in the interview of, for example, a burglary suspect; the use of brutal language or detail may only be suspicious when used in relation to a close relative, and not, for example, in a witness statement about a crime involving a stranger.

The next task in the present thesis was to conduct an empirical investigation into the cues suggested by the participants, and establish whether cues of this kind have any utility in discriminating between honest and deceptive appeals.



# **Chapter 7**

## **Studies 2 and 3: Quantitative Investigation of Cues to Deception Generated in Study 1**

### **7.1. Introduction**

The aim of Study 1 was to use a qualitative methodology to generate potential cues for further examination; in this chapter, two further studies are described (Studies 2 and 3), which represent a more systematic, two-part empirical investigation of the cues thus identified, using the same set of stimulus materials. Study 2 is an investigation of cues to deception using subjective observer ratings of individuals making appeals. As discussed in Chapter 2, although meta-analytic findings suggest that honesty and deception may be discriminated more powerfully by using subjective measures rather than objective measures (DePaulo et al., 2003), studies investigating subjective measures of high stakes deception are relatively scarce (see, for example, Koper & Sahlman, 1991; Mann & Vrij, 2006; Vrij & Mann, 2001a).

Study 3 is an investigation of verbal cues to deception using more objective frequency counts on transcripts of the appeals. It was decided to investigate verbal cues with frequency counts as well as with subjective observer ratings, as it was

recognised that some verbal cues may be difficult for observers to judge. Moreover, in applied contexts frequency counts of behaviours may sometimes be more useful as they are more easily verified.

The first aim of the present two studies, therefore, was to investigate which of the cues reported by the accurate participants in Study 1 were useful in discriminating between honest and deceptive behaviour on the same set of appeals (the intention being to identify cues to test on a larger sample in a future study). As discussed in Chapter 1, several researchers have suggested that the more cues to deception that a person exhibits, the more likely that person is to be lying, and therefore it may be useful to look for clusters of behaviour (for example, DePaulo et al., 2003; ten Brinke & Porter, 2012; Vrij, 2004). Consequently, a further aim of the two studies in this chapter was to utilise a case by case methodology, to assess the potential for the classification of individual appeals as honest or deceptive, based on the cues that were found to discriminate between honest and deceptive appeals. As there are almost certainly no universal cues to deception, the use of a range of cues to assess individual cases overcomes the evident problems of focusing on a single cue to determine veracity. Previous studies (Ekman et al., 1991; Leal et al., 2010) have found high accuracy rates using the case by case method of discrimination by subject, rather than looking at differences between groups (86% and 79% respectively). It was, therefore, hypothesised that classifying subjects as honest or deceptive by considering the results across cues, would achieve high overall accuracy in discriminating between honest and deceptive appealers.

Finally, as noted in Chapter 3, studies using real life, high stakes lies as stimulus materials have found observer accuracy in detecting deception to be above chance (Mann & Vrij, 2006; Mann et al., 2004, 2006; Vrij et al., 2006); however, as

the observers used in these studies were police officers, it is not clear whether it was the ecological validity of the stimulus materials, or the biased sample of observers, that produced the high accuracy rates. Consequently, in Study 2, observer accuracy in detecting deception in a non-police sample was investigated; it was hypothesised that the observers would achieve accuracy rates above chance, due to the use of high stakes stimulus materials. (As Study 3 involved use of two coders to count frequencies of verbal cues in transcripts it was not included in the accuracy investigation.)

As part of the investigation of the accuracy of observers in detecting deception, the data were also analysed for the presence of a truth bias. As discussed in Chapter 1, truth bias refers to the phenomenon that more truth-tellers than liars are correctly identified, because, overall, there is a bias towards judging subjects to be truthful. This is well documented in the deception literature (for example, Bond & DePaulo, 2006; Levine, Park & McCornack, 2009). However, it was hypothesised that the use of ecologically valid stimulus materials might eliminate truth bias. The rationale for this is that, in contexts where cues to deception are weak and difficult to discern, people may be more likely to rely on a truthfulness heuristic. However, it may be the case that high stakes lies produce cues to deception that are easier for observers to discern, precluding the assumption of honesty.

## **7.2. STUDY 2: Subjective Observer Ratings of Cues to Deception**

To reiterate, Study 2 was an empirical investigation to validate the cues to deception and to honesty identified in Study 1, using the same materials but with a different,

independent panel of observers. It was expected that utilising case by case analyses to classify subjects as honest or deceptive by considering the results across cues, would achieve high overall accuracy in discriminating between liars and truth-tellers. It was also hypothesised that the use of ecologically valid stimulus materials would result in above chance accuracy in detecting deception, and the elimination of truth bias.

### ***7.2.1. Subjective Observer Ratings: Method***

#### *7.2.1.1. Participants*

Five adult observers (two male and three female), were used to rate the materials. Ages of the observers ranged from 21 to 75 ( $M = 42.00$ ;  $SD = 19.00$ ). One observer was a postgraduate psychology student and four were non-psychologists. All were recruited to take part on a voluntary basis, using opportunity sampling, and none had previously received any formal training in lie detection techniques. It is usual in deception research to use as few as two or three people to rate behaviours (see, for example, Davis et al., 2005; Harpster et al., 2009; Koper & Sahlman, 1991; Mann et al., 2002; ten Brinke & Porter, 2012; Villar et al., 2011; Vrij & Mann, 2001a), as it is normally assumed that, if the ratings were not reliable, they would be unlikely to statistically discriminate between liars and truth-tellers (i.e. unreliable ratings are more likely to give rise to Type II errors).

#### *7.2.1.2. Materials and procedure*

The appeals were the same as those used in Study 1. Observers participated individually, and attended a psychology laboratory at the University of Liverpool. Participation lasted approximately one hour. Observers were shown an appeal, and were then required to complete a response sheet for each appeal. The response sheet contained a brief outline of the content of the appeal, for example, ‘This clip features a man appealing in the murder of his wife.’ Observers were asked to indicate whether they were familiar with the individual featured in the clip. The response sheet contained one statement relating to each of the 50 cues identified in Study 1 (for example, ‘This person’s emotions are fake’, and ‘This person is focused on the missing/murdered person’). Observers were asked to indicate their agreement with each statement on a five point Likert scale from 1 (strongly disagree) to 5 (strongly agree). There were two statements relating to hope (‘This person is hopeful of finding the missing person’ and ‘This person does not seem to have any hope of finding the missing person’), and participants were instructed to check a N/A box instead of rate the behaviour in cases where a body had already been found (i.e. the relative was publicly known to be dead, not missing). At the end of each response sheet, observers were asked to indicate whether they thought the appellant was deceptive or honest. See Appendix 4 for the full response sheet.

This process was repeated for each appeal. The appeals were presented in a different, randomised order for each observer, and each observer completed the rating procedure individually.

### **7.2.2. Subjective Observer Ratings: Results and Discussion**

#### *7.2.2.1. Inter-rater reliability*

To assess inter-rater reliability between the five observers, Kendall's W was calculated for each cue. Inter-rater reliability was not significant ( $p > .05$ ) for ten of the cues, which were dropped from further analysis (calm and calculated, lack of emotion in voice, use of first person pronouns, irrelevant/generic statements, trying to be sensible, avoid brutal language/detail, positive emotion expressed towards relative, smiles appropriate to speech content, focused on the relative, and behaving as expected). See Table 7.1.

#### *7.2.2.2. Analyses of subjective cues rated by observers*

The mean score for each cue in each appeal was calculated from the scores of the five observers, so that each appeal had a single score for each cue; these scores were used in all further analyses. One observer was familiar with the outcome of one of the nine appeals, and as ratings of the cues were required to be independent of the observers' knowledge of the veracity of the appeals, ratings by this observer for this one appeal were replaced with the mean rating of the other observers.

To investigate which cues were associated with deception, a series of one way ANOVAs was conducted on the means for each cue between the deceptive appealers and the honest appealers (veracity condition). Notwithstanding the small sample sizes (five and four), parametric ANOVAs were used; these were considered appropriate for two reasons. First, the main purpose of this study was exploratory, to identify cues worthy of further investigation; hence a statistically powerful test was considered most useful. Second, the parametric ANOVA is fairly robust to violations

of the standard assumptions with fixed levels of the independent variable and fairly equal sample sizes (Glass & Stanley, 1970). A series of ANOVAs was also chosen rather than a MANOVA, as there were more dependent variables than cases per cell, thus violating the minimal sample size requirement for MANOVA (Tabachnick & Fidell, 2001).

As shown in Table 7.1, cues which yielded significant effects between deceptive and honest appealers were as follows: fake emotion, fake facial expression, putting on a performance, high vocal pitch, lists, gaze aversion, lack of sympathy for the appealers, and creepy, were related to deceptive appeals. Sad eyes, genuine facial expression, and containing emotion, were related to honest appeals. Two cues only very narrowly missed significance on a two-tailed test ( $p < .06$ ), and had means in the expected directions; these were, genuinely sad and credible eye contact. Again, in view of the exploratory nature of the study, no correction for multiple testing was applied.

Since the sample size was small, and consequently there was a risk of Type II errors in analyses of statistical significance, it was considered useful to examine effect sizes as well, in order to illuminate meaningful differences between deceptive and honest appeals (Howell, 2002). According to Cohen (1988), effect sizes of 0.8 can be considered as large, and a number of cues showed effect sizes of this magnitude ( $d \geq 0.80$ ) in the predicted directions. These were, pretending to cry, no emotional variation, emotionally cold, focusing on others, equivocation, does not make sense, defensive, behaving unnaturally, and disliking the appealers, all of which were associated with deceptive appeals. Calmness appropriate to speech content, genuine/heartfelt appeal, genuine behaviours, normal, feeling sorry for the appealers, and feeling the appealers' pain, were associated with honest appeals.

In sum, 13 cues were found to discriminate significantly, or very near significantly, between honest and deceptive appealers, with a further 15 cues having substantive effect sizes in the expected direction. Indeed, of the 38 cues included in the analyses, means were in the expected direction for 34 cues, a result which was statistically significant by binomial test ( $p < .0001$ ). Furthermore, 27 cues had large effect sizes ( $d \geq 0.80$ ) in the expected direction, also a significant deviation from chance ( $p = .004$ ). These findings suggest that a large number of behaviours identified by the accurate participants in Study 1 may potentially have some utility in discriminating between deceptive and honest appeals.

It was not possible to run ANOVAs on the two cues ‘hopeful’ and ‘lack of hope’, as the cues were rated on only two lying appealers and two honest appealers (the cases in which no body had been found). Therefore, an alternative analysis was performed: for each observer, the mean of the ‘hopeful’ and ‘lack of hope’ ratings (on a five point scale) given to the two lying appealers, and the mean given to the two honest appealers, was calculated. A Wilcoxon test was then performed on these data, to compare the mean ‘hopeful’ and ‘lack of hope’ ratings given to lying appealers and honest appealers. As expected, honest appealers (median score 4) were rated significantly higher than deceptive appealers (median score 2) on the cue ‘hopeful’ ( $Z = 1.84$ ,  $p = .033$  one-tailed). Also as expected, deceptive appealers (median score 4) were rated as significantly higher than honest appealers (median score 2) on the cue ‘lack of hope’ ( $Z = 1.89$ ,  $p = .030$  one-tailed).



Table 7.1: *Kendall's W for all cues rated subjectively by observers; means, SDs, CIs, F ratios and p values, and effect sizes for cues which achieved significant inter-rater reliability*

Cue	Kendall's W (p)	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	F ratio (p)	Effect size
<b><u>Cues to deception</u></b>					
Fake emotion	.61 (p = .002)	<b>3.66</b> (0.32) 3.15 – 4.17	<b>2.48</b> (0.84) 1.43 – 3.53	F (1,7) = 6.90 p = .034	d = 1.86
Fake facial expression	.44 (p = .024)	<b>3.46</b> (0.24) 3.09 – 3.84	<b>2.56</b> (0.64) 1.77 – 3.35	F (1,7) = 7.04 p = .033	d = 1.86
Performance	.61 (p = .002)	<b>3.98</b> (0.46) 3.24 – 4.71	<b>2.72</b> ( 0.91) 1.59 – 3.85	F (1,7) = 6.16 p = .042	d = 1.73
Pretending to cry	.73 (p < .001)	3.16 (1.16) 1.32 – 5.01	2.04 (0.61) 1.29 – 2.79	F (1,7) = 3.56 p = .101	d = 1.21
High vocal pitch	.53 (p = .007)	<b>3.25</b> (0.91) 1.79 – 4.71	<b>2.08</b> (0.30) 1.70 – 2.46	F (1,7) = 7.40 p = .030	d = 1.73
No emotional variation	.43 (p = .028)	3.63 (0.42) 2.96 – 4.29	3.00 (0.68) 2.16 – 3.84	F (1,7) = 2.57 p = .153	d = 1.10
Lack of emotion	.51 (p = .009)	2.83 (0.69) 1.72 – 3.93	3.04 (1.03) 1.76 – 4.32	F (1,7) = 0.13 p = .733	d = 0.25
Painless recall	.52 (p = .007)	2.80 (0.63) 1.79 – 3.81	2.92 (1.03) 1.65 – 4.19	F (1,7) = 0.04 p = .845	d = 0.14
Calm / calculated	.31 (p = .143)				
Lack emotion in voice	.31 (p = .128)				
Cold	.43 (p = .027)	3.54 (0.37) 2.95 – 4.12	2.60 ( 0.97) 1.40 – 3.80	F (1,7) = 3.28 p = .113	d = 1.28
Lack of facial expression	.39 (p = .049)	3.33 (0.63) 2.32 – 4.33	2.96 (0.83) 1.93 – 3.99	F (1,7) = 0.53 p = .492	d = 0.49
Brutal lang./detail	.50 (p = .011)	1.93 (0.85) 0.57 – 3.28	1.76 (0.55) 1.07 – 2.43	F (1,7) = 0.12 p = .735	d = 0.22
Impersonal	.40 (p = .045)	3.20 (0.40) 2.56 – 3.84	2.76 (0.99) 1.53 – 3.99	F (1,7) = 0.68 p = .437	d = 0.58
Focus on others	.57 (p = .004)	3.45 (0.53) 2.61 – 4.29	2.72 (0.94) 1.55 – 3.89	F (1,7) = 1.89 p = .212	d = 0.96
First person pronouns	.30 (p = .153)				

Cue	Kendall's W ( <i>p</i> )	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
Equivocation	.42 ( <i>p</i> = .031)	2.79 (0.17) 2.52 – 3.05	2.20 (0.57) 1.50 – 2.90	<i>F</i> (1,7) = 3.94 <i>p</i> = .087	<i>d</i> = 1.40
Irrelevant/gen eric statement	.35 ( <i>p</i> = .084)				
Lists	.48 ( <i>p</i> = .013)	<b>3.71</b> (0.51) 2.90 – 4.53	<b>2.56</b> (0.78) 1.59 – 3.53	<i>F</i> (1,7) = 6.41 <i>p</i> = .039	<i>d</i> = 1.75
Gaze aversion	.65 ( <i>p</i> = .001)	<b>4.13</b> (0.49) 3.35 – 4.90	<b>2.60</b> (0.93) 1.45 – 3.75	<i>F</i> (1,7) = 8.72 <i>p</i> = .021	<i>d</i> = 2.04
Does not make sense	.61 ( <i>p</i> = .002)	3.24 (0.46) 2.50 – 3.97	2.40 (0.93) 1.25 – 3.55	<i>F</i> (1,7) = 2.68 <i>p</i> = .146	<i>d</i> = 1.14
Making assumptions	.49 ( <i>p</i> = .012)	2.53 (0.66) 1.47 – 3.58	2.80 (0.73) 1.89 – 3.71	<i>F</i> (1,7) = 0.40 <i>p</i> = .579	<i>d</i> = 0.40
Lack of hope*		<b>3.40</b> (0.28)	<b>2.40</b> (0.00)		<i>d</i> = 1.42
Indication of dispute	.53 ( <i>p</i> = .007)	2.25 (0.30) 1.77 – 2.73	2.16 (0.71) 1.28 – 3.05	<i>F</i> (1,7) = 0.06 <i>p</i> = .822	<i>d</i> = 0.17
Defensive	.46 ( <i>p</i> = .019)	2.89 (0.27) 2.46 – 3.31	2.20 (0.63) 1.41 – 2.99	<i>F</i> (1,7) = 4.06 <i>p</i> = .084	<i>d</i> = 1.42
Fear/nervousn ess	.51 ( <i>p</i> = .009)	2.85 (1.10) 1.10 – 4.60	2.64 (0.46) 2.07 – 3.21	<i>F</i> (1,7) = 0.15 <i>p</i> = .707	<i>d</i> = 0.25
Dislike	.48 ( <i>p</i> = .013)	3.49 (0.36) 2.91 – 4.06	2.64 (0.97) 2.36 – 3.67	<i>F</i> (1,7) = 2.67 <i>p</i> = .146	<i>d</i> = 1.16
Creepy	.41 ( <i>p</i> = .040)	<b>3.54</b> (0.37) 2.95 – 4.12	<b>2.24</b> (0.74) 1.32 – 3.16	<i>F</i> (1,7) = 10.08 <i>p</i> = .016	<i>d</i> = 2.22
Behaving unnaturally	.61 ( <i>p</i> = .002)	3.63 (0.62) 2.63 – 4.62	2.88 (1.09) 1.52 – 4.24	<i>F</i> (1,7) = 1.46 <i>p</i> = .267	<i>d</i> = 0.83
No sympathy	.58 ( <i>p</i> = .003)	<b>3.71</b> (0.36) 3.14 – 4.29	<b>2.48</b> (0.92) 1.33 – 3.63	<i>F</i> (1,7) = 6.22 <i>p</i> = .041	<i>d</i> = 1.76

**Cues to  
honesty**

Genuinely sad	.61 ( <i>p</i> = .002)	<b>2.4</b> (0.49) 1.62 – 3.18	<b>3.4</b> (0.76) 2.45 – 4.35	<i>F</i> (1,7) = 5.12 <i>p</i> = .058	<i>d</i> = 1.56
Facial expression	.51 ( <i>p</i> = .009)	<b>2.49</b> (0.24) 2.11 – 2.87	<b>3.44</b> (0.75) 2.50 – 4.38	<i>F</i> (1,7) = 5.78 <i>p</i> = .047	<i>d</i> = 1.71
Sad eyes	.58 ( <i>p</i> = .003)	<b>2.19</b> (0.41) 1.54 – 2.84	<b>3.22</b> (0.85) 2.38 – 4.50	<i>F</i> (1,7) = 7.15 <i>p</i> = .032	<i>d</i> = 1.54

Cue	Kendall's W ( <i>p</i> )	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
Desperation/ Urgency	.53 ( <i>p</i> = .007)	2.66 (0.91) 1.22 – 4.11	2.64 ( 0.74) 1.72 – 3.56	<i>F</i> (1,7) = 0.002 <i>p</i> = .968	<i>d</i> = 0.02
Containing emotion	.43 ( <i>p</i> = .030)	<b>2.40</b> (0.59) 1.46 – 3.34	<b>3.60</b> (0.42) 3.07 – 4.13	<i>F</i> (1,7) = 12.73 <i>p</i> = .009	<i>d</i> = 2.34
Calm approp. to speech	.59 ( <i>p</i> = .003)	2.55 (0.66) 1.50 – 3.60	3.32 (0.50) 2.70 – 3.94	<i>F</i> (1,7) = 3.98 <i>p</i> = .086	<i>d</i> = 1.32
Trying to be sensible	.37 ( <i>p</i> = .063)				
Hopeful*		<b>2.40</b> (0.00)	<b>3.70</b> (0.14)		<i>d</i> = 13.13
Avoid brutal lang./detail	.21 ( <i>p</i> = .395)				
Genuine/heart felt appeal	.58 ( <i>p</i> = .003)	2.29 (0.56) 1.39 – 3.18	3.24 (1.06) 1.92 – 4.56	<i>F</i> (1,7) = 2.58 <i>p</i> = .152	<i>d</i> = 1.12
No prepared speech	.60 ( <i>p</i> = .002)	3.43 (0.85) 2.07 – 4.78	2.80 ( 1.05) 1.50 – 4.10	<i>F</i> (1,7) = 0.93 <i>p</i> = .368	<i>d</i> = 0.66
Pos. emotion to relative	.32 ( <i>p</i> = .123)				
Smile approp. to speech	.21 ( <i>p</i> = .398)				
Genuine behaviours	.57 ( <i>p</i> = .004)	2.25 (0.57) 1.34 – 3.16	3.24 (0.80) 2.24 – 4.24	<i>F</i> (1,7) = 4.26 <i>p</i> = .078	<i>d</i> = 1.43
Credible eye contact	.61 ( <i>p</i> = .002)	<b>2.19</b> (0.41) 1.54 – 2.84	<b>3.32</b> (0.89) 2.22 – 4.43	<i>F</i> (1,7) = 5.44 <i>p</i> = .052	<i>d</i> = 1.63
Focused on relative	.36 ( <i>p</i> = .076)				
Feel appealer pain	.51 ( <i>p</i> = .010)	2.49 (0.33) 1.96 – 3.02	3.20 (0.91) 2.08 – 4.32	<i>F</i> (1,7) = 2.19 <i>p</i> = .183	<i>d</i> = 1.04
Feel sorry for appealer	.53 ( <i>p</i> = .007)	2.49 ( 0.24) 2.11 – 2.87	3.32 (0.88) 2.23 – 4.41	<i>F</i> (1,7) = 3.31 <i>p</i> = .112	<i>d</i> = 1.29
Normal	.50 ( <i>p</i> = .011)	2.23 (0.39) 1.61 – 2.84	3.04 (0.92) 1.90 – 4.18	<i>F</i> (1,7) = 2.691 <i>p</i> = .145	<i>d</i> = 1.15
Behaving as expected	.36 ( <i>p</i> = .077)				

Means in bold differ significantly/marginally significantly from each other. \* Wilcoxon test performed on mean of observer ratings for appeals in which relative was reportedly missing

#### 7.2.2.3. *Case by case analyses using subjective cues rated by observers*

As an additional way of construing the same data, case by case analyses were also conducted to assess the potential of each cue found to be significant to predict veracity. This idea has been used previously in the literature (see, for example, Ekman et al., 1991; Leal et al., 2010), and has been considered important on the grounds that, even if we were to assume that any significant results are robust and reliable in a statistical sense, it would not necessarily follow that they would have any practical predictive value in the field. In other words, the idea here is not to assess whether these cues significantly predict veracity (as by the criterion for selection, inevitably they must), but the *relative extent* to which they might predict veracity, assuming they are reliable.

The cues considered in this analysis were those that had significantly discriminated between liars and truth-tellers, and also the cues ‘hopeful’ and ‘no hope’ which, although could not be demonstrated to discriminate significantly between liars and truth-tellers, observers were found to rate liars and truth-tellers significantly differently on these two cues, in the directions expected. For each case (appeal), the score for each cue was compared with the overall cue mean. For the cues related to honesty, scores that fell above the cue mean for the honest appeals were classified as a ‘hit’, and scores that fell below the cue mean for deceptive appeals were classified as a ‘hit’. For the cues related to deception, scores that fell above the cue mean for the deceptive appeals were classified as a ‘hit’, and scores that fell below the cue mean for honest appeals were classified as a ‘hit’. Scores that were equal to the overall cue mean were considered unclassified and not included in the hit/miss accuracy calculations. In this way, a percentage of ‘hits’ was calculated

for each cue (i.e. cases combined). The cues had 'hit' rates between 78% and 100% and a mean hit rate of 88%. See Table 7.3.

The information from the cues was then combined to classify individual cases as deceptive or honest. For each case, the number of 'hits' on the cues was calculated (i.e. the number of times the appellant scored in the expected direction on each of the cues). Honest appeals had 'hit' rates between 38 - 100%, deceptive appeals had 'hit' rates between 91 - 100%, and the mean 'hit' rate was 85%. See Table 7.3. Cases which scored above 50% 'hits' on the cues were categorised as correctly classified (the appellant scored in the expected direction on the majority of the cues), and cases that scored below 50% 'hits' on the cues were categorised as misclassified (the appellant did not score in the expected direction on the majority of the cues). In this way, all of the four deceptive appeals were correctly classified, and three of the five honest appeals were correctly classified. Overall, therefore, using these criteria, 78% of cases were correctly classified, supporting the hypothesis that using a combination of cues would result in high classification accuracy.

#### *7.2.2.4. Accuracy of observers' credibility judgements*

Individual observer accuracy ranged from 56% to 100%, and mean accuracy was 75%. To test the hypothesis that observers would be able to discriminate between liars and truth-tellers above the level of chance, for each observer, the percentage of appeals correctly classified on veracity was compared with the percentage of appeals incorrectly judged on veracity, using a Wilcoxon test. Observers correctly classified appeals (median percentage = 67) significantly more often than they incorrectly classified appeals (median percentage = 33) ( $Z = 2.03$ ,  $p = .042$ , two-tailed).

Analyses were then conducted to investigate the hypothesis that observers would not demonstrate a truth bias. For each observer, the percentage of deceptive appeals correctly classified was compared with the percentage of deceptive appeals not correctly classified using a Wilcoxon test. Observers correctly classified deceptive appeals (median percentage = 100) significantly more often than they incorrectly classified deceptive appeals (median percentage = 0), ( $Z = 2.07$ ,  $p = .038$ , two-tailed), suggesting that observers were able to accurately classify deceptive appeals. Again, for each observer, the percentage of honest appeals correctly classified was compared with the percentage of honest appeals not correctly classified using a Wilcoxon test. There was no significant difference ( $Z = 0.71$ ,  $p = .480$ , two-tailed) between observers' correct classifications of honest appeals (median percentage = 60) and their incorrect classifications of honest appeals (median percentage = 40), suggesting that observers were not able to accurately classify honest appeals. Finally, the percentage of deceptive appeals correctly classified was compared with the percentage of honest appeals correctly classified using a Wilcoxon test. The difference between observers' correct classifications of deceptive appeals (median percentage = 100) and honest appeals (median percentage = 60), approached significance ( $Z = 1.83$ ,  $p = .068$ , two-tailed), suggesting that observers were somewhat more able to correctly classify liars than truth-tellers, the opposite pattern of a truth bias.

### **7.2.3. Summary of Study 2**

The primary aim of Study 2 was to investigate which of the cues reported by the accurate participants Study 1 were useful in discriminating between honest and deceptive appeals. Thirteen cues were found to discriminate significantly/marginally significantly between honest and deceptive appealers, with a further 14 cues having large effect sizes in the expected direction, suggesting that the behaviours identified by the accurate participants in Study 1 warrant further investigation. As hypothesised, the case by case analyses yielded high accuracy rates, suggesting that combining cues to investigate clusters of behaviour may be a productive approach in discriminating between liars and truth-tellers. Moreover, it appears that the use of ecologically valid stimulus materials may have an effect on how accurately people are able to detect deception. As hypothesised, observers were able to discriminate between honest and deceptive appealers at a rate well above chance, and truth bias was eliminated.

Having investigated subjective observer ratings of cues present in the video materials, attention was turned to frequency counts of verbal cues in transcripts of the appeals.

### **7.3. STUDY 3: Frequency Counts of Verbal Cues to Deception**

To reiterate, the primary aim of this next study, Study 3, was to investigate which of the verbal cues reported by the accurate participants in Study 1 were useful in discriminating between honest and deceptive appeals. It was also expected that case

by case analyses, that classify subjects as honest or deceptive by considering the results across cues, would achieve high overall accuracy in discriminating between liars and truth-tellers.

### **7.3.1. Verbal Frequency Counts: Method**

#### *7.3.1.1. Materials and Procedure*

The audio recordings of the nine appeals used in Study 2 were transcribed. Transcripts varied in length between 79 and 229 words. See Appendix 5 for transcripts of the appeals. Two coders then counted frequencies of predetermined verbal behaviours in the transcripts of the appeals. Both coders were graduate students studying deception. One coder knew which appeals were deceptive and which were honest, and one coder was blind to the veracity of the appeals. Inter-rater reliability for the frequency counts was high, ranging from  $r = .71$  to  $r = 1.00$ , with a mean of  $r = .91$  (see Table 7.2). The verbal behaviours included in the study were based on those identified in Study 1 as used by people accurate at detecting deception, plus an additional three cues which have been found to discriminate between honesty and deception across contexts in previous research on high stakes deception.

Verbal behaviours identified in Study 1 that were expected to be related to deception in the context of an appeal were: using brutal language or detail about the relative (for example, ‘murdered’ or ‘blood’); focus on others (measured by references to other people or groups of people, not including the police or the perpetrator); deletion of first person singular pronouns (pronouns disappearing



altogether, for example ‘Just want my wife back’, or replacing/switching first person with second/third person pronouns, for example ‘and I, eh, what do you do?’); equivocation (modifiers, minimisers and hedging words, for example, ‘kind of’ ‘a little’ ‘probably’); sentences or statements that are irrelevant or generic; lists of three or more items (for example, of people, actions, emotions); sentences or statements that do not make sense (for example, are illogical or unclear); and expressions of lack of hope (in cases in which no body had been found).

Also included were three verbal cues previously found to discriminate between honesty and deception in studies using real life, high stakes lies: pauses, speech disturbances, and repetition. Pauses have been associated with deception in two studies using police suspect interviews (Vrij & Mann, 2001a; Mann et al., 2002), and with deception in suspect statements (Davis et al., 2005). In the present study separate frequency counts were included for filled pauses, unfilled pauses, and pauses in the middle of a clause. Speech disturbances/dysfluency have also been associated with deception in suspect statements (Davis et al., 2005), and in police interviews (Vrij & Mann, 2001a), so speech errors were included here (for example, incomplete sentences, grammatical errors, and single word repetition). Repetition of a phrase has been found to discriminate between honesty and deception in 911 homicide calls (Harpster et al., 2009), and in suspect statements (Davis et al., 2005). (See Appendix 4 for coders’ full instructions for frequency counts.)

Verbal behaviours identified in Study 1 that were expected to relate to honesty in the context of an appeal were: expressions of hope (relevant in cases in which no body had been found); avoiding the use of brutal language (for example, saying ‘taken from us’ instead of ‘murdered’); positive emotion expressed towards the relative (including positive descriptions of the relative, terms of endearment, and

expressions of love for the relative); focusing on the relative (references to the relative by name, in the second person, in the third person, and in terms of their relationship to the appellant); and use of first person pronouns (references to first person singular and first person plural).

### ***7.3.2. Verbal Frequency Counts: Results and Discussion***

#### *7.3.2.1. Analyses of verbal cues using frequency counts*

For each verbal cue, a mean frequency count was calculated using the data from the two coders. Preliminary analyses showed no significant results for the length of the transcript, and so this was not included in further analysis. Nevertheless, to standardise the frequency counts, a score was calculated of the mean frequency counts as a percentage of the total length of the transcript, and this score was used in all further analyses.

To investigate which cues were associated with deception, a series of one way ANOVAs was conducted for each cue between the four liars and the five truth-tellers (veracity condition). Again, a series of ANOVAs was also chosen rather than a MANOVA, as there were more dependent variables than cases per cell, thus violating the minimal sample size requirement for MANOVA (Tabachnick & Fidell, 2001). Means, standard deviations, confidence intervals, *F* ratios, *p* values and effect sizes are presented in Table 7.2.

ANOVAs for cues which yielded significant effects for veracity are as follows: focus on others; deletion of first person pronouns; statements or sentences that do not make sense; repetition; and positive emotion expressed towards the

relative. For all significant cues, means were in the expected direction, i.e. for the cue expected to be associated with honesty (positive emotion expressed towards the relative), honest appealers showed higher frequencies than deceptive appealers; and for the cues expected to be associated with deception (focusing on others; deletion of the first person pronouns; statements or sentences that do not make sense; and repetition of a phrase), deceptive appealers showed higher frequencies than honest appealers; see Table 7.2.

As in Study 2, it was not possible to run ANOVAs on the two cues ‘hopeful’ and ‘no hope’, as the cues were coded on only two deceptive appealers and two honest appealers (the cases in which no body had been found). However, means were in the expected direction, with honest appealers expressing more hope ( $M = 2.80$ ,  $SD = 1.41$ ) than lying appealers ( $M = 0.00$ ,  $SD = 0.00$ ), and with lying appealers expressing lack of hope ( $M = 0.30$ ,  $SD = 0.42$ ) more than honest appealers ( $M = 0.00$ ,  $SD = 0.00$ ).

Again, due to the small sample size, effect size analyses were used to illuminate non-significant trends in the data (see Table 7.2). There were large differences ( $d \geq 0.80$ ), all in the expected direction, between deceptive and honest appealers in the following cues: using brutal language or detail, equivocation, irrelevant/generic statements, lists, lack of hope, and filled pauses ( $d = 0.79$ ), had a higher frequency in deceptive appeals. Expressions of hope, and avoiding brutal language or detail, had a higher frequency in honest appeals.

Table 7.2: *Inter-rater reliability, means, (SDs), CIs, F ratios and p values, and effect sizes for frequency counts of each verbal cue*

Cue	Inter-rater reliability ( <i>r</i> )	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
<b><u>Cues to deception</u></b>					
Use brutal language/detail	<i>r</i> = .71 <i>p</i> = .033	0.44 (0.54) -0.42 – 1.29	0.08 (0.19) -0.15 – 0.32	<i>F</i> (1,7) = 1.94 <i>p</i> = .207	<i>d</i> = 0.89
Focus on others	<i>r</i> = .93 <i>p</i> < .001	<b>2.95 (2.63)</b> -1.23 – 7.13	<b>0.20 (0.29)</b> -0.15 – 0.56	<i>F</i> (1,7) = 5.57 <i>p</i> = .050	<i>d</i> = 1.47
Deletion of 1 <sup>st</sup> person	<i>r</i> = .95 <i>p</i> < .001	<b>0.90 (0.74)</b> -0.30 – 2.10	<b>0.00 (0.00)</b> 0.00 – 0.00	<i>F</i> (1,7) = 7.42 <i>p</i> = .030	<i>d</i> = 1.72
Equivocation	<i>r</i> = 1.00 <i>p</i> < .001	3.12 (1.55) 0.66 – 5.58	1.23 (1.73) -0.91 – 3.37	<i>F</i> (1,7) = 2.93 <i>p</i> = .131	<i>d</i> = 1.15
Irrelevant/generic statements	<i>r</i> = .95 <i>p</i> < .001	0.60 (0.74) -0.58 – 1.79	0.09 (0.19) -0.15 – 0.33	<i>F</i> (1,7) = 2.30 <i>p</i> = .173	<i>d</i> = 0.94
Lists	<i>r</i> = .89 <i>p</i> = .001	1.18 (0.81) -0.11 – 2.48	0.41 (0.49) -0.20 – 1.02	<i>F</i> (1,7) = 3.14 <i>p</i> = .120	<i>d</i> = 1.15
Does not make sense	<i>r</i> = .88 <i>p</i> = .002	<b>1.10 (0.30)</b> 0.62 – 1.58	<b>0.09 (0.19)</b> -0.15 – 0.33	<i>F</i> (1,7) = 37.75 <i>p</i> < .001	<i>d</i> = 4.02
Lack of hope	<i>r</i> = 1.00 <i>p</i> < .001	0.30 (0.42)	0.00 (0.00)		<i>d</i> = 1.01
Filled pauses	<i>r</i> = .92 <i>p</i> < .001	1.16 (0.83) -0.17 – 2.48	0.58 (0.62) -0.19 – 1.35	<i>F</i> (1,7) = 1.42 <i>p</i> = .273	<i>d</i> = 0.79
Unfilled pauses	<i>r</i> = 1.00 <i>p</i> < .001	12.50 (4.34) 5.60 – 19.40	13.61 (3.63) 9.10 – 18.12	<i>F</i> (1,7) = 0.18 <i>p</i> = .687	<i>d</i> = 0.23
Pauses in a clause	<i>r</i> = .87 <i>p</i> = .003	3.65 (1.84) 0.73 – 6.57	2.67 (2.40) -0.31 – 5.65	<i>F</i> (1,7) = 0.45 <i>p</i> = .525	<i>d</i> = 0.46
Speech errors	<i>r</i> = .71 <i>p</i> = .031	2.11 (0.68) 1.02 – 3.20	1.47 (2.73) -1.91 – 4.86	<i>F</i> (1,7) = 0.20 <i>p</i> = .667	<i>d</i> = 0.32
Repetition	<i>r</i> = .86 <i>p</i> = .003	<b>1.76 (0.36)</b> 1.19 – 2.34	<b>0.70 (0.53)</b> 0.04 – 1.36	<i>F</i> (1,7) = 11.46 <i>p</i> = .012	<i>d</i> = 2.34
<b><u>Cues to honesty</u></b>					
Hope	<i>r</i> = .94 <i>p</i> = .057	0.00 (0.00)	2.80 (1.41)		<i>d</i> = 2.82
Avoids brutal language/detail	<i>r</i> = 1.00 <i>p</i> < .001	0.00 (0.00) 0.00 – 0.00	0.75 (0.76) -0.19 – 1.69	<i>F</i> (1,7) = 3.86 <i>p</i> = .090	<i>d</i> = 1.41
Pos. emotion to relative	<i>r</i> = .99 <i>p</i> < .001	<b>0.41 (0.56)</b> -0.48 – 1.30	<b>4.63 (2.65)</b> 1.35 – 7.92	<i>F</i> (1,7) = 9.60 <i>p</i> = .017	<i>d</i> = 2.20

Cue	Inter-rater reliability ( <i>r</i> )	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
Refer to relative by name	<i>r</i> = 1.00 <i>p</i> < .001	0.95 (0.96) -0.58 – 2.49	1.90 (1.78) -0.31 – 4.10	<i>F</i> (1,7) = 0.90 <i>p</i> = .374	<i>d</i> = 0.66
Refer to relative in 2 <sup>nd</sup> person	<i>r</i> = .99 <i>p</i> < .001	1.92 (3.85) -4.20 – 8.04	0.48 (1.08) -0.86 – 1.82	<i>F</i> (1,7) = 0.66 <i>p</i> = .444	<i>d</i> = 0.51
Refer to relative in 3 <sup>rd</sup> person	<i>r</i> = .71 <i>p</i> = .034	2.15 (1.93) -0.92 – 5.23	2.92 (1.95) 0.50 – 5.33	<i>F</i> (1,7) = 0.35 <i>p</i> = .576	<i>d</i> = 0.40
Refer to relative by relationship	<i>r</i> = .94 <i>p</i> < .001	1.34 (0.94) -0.15 – 2.83	1.33 (1.22) -0.18 – 2.85	<i>F</i> (1,7) = 0.00 <i>p</i> = .989	<i>d</i> = 0.01
Total references to relative*		6.37 (3.94) 0.11 – 12.64	6.63 (3.74) 1.98 – 11.28	<i>F</i> (1,7) = 0.01 <i>p</i> = .923	<i>d</i> = 0.01
Use 1 <sup>st</sup> person singular	<i>r</i> = .98 <i>p</i> < .001	5.89 (3.52) 0.28 – 11.49	3.46 (3.90) -1.38 – 8.30	<i>F</i> (1,7) = 0.94 <i>p</i> = .366	<i>d</i> = 0.65
Use 1 <sup>st</sup> person plural	<i>r</i> = .77 <i>p</i> = .015	2.57 (1.93) -0.51 – 5.64	3.45 (3.34) -0.69 – 7.60	<i>F</i> (1,7) = .22 <i>p</i> = .654	<i>d</i> = 0.32
Total self references*		8.45 (3.62) 2.69 – 14.22	6.91 (3.13) 3.02 – 10.80	<i>F</i> (1,7) = 0.47 <i>p</i> = .515	<i>d</i> = 0.46

Means in bold differ significantly/marginally significantly from each other. \* Cue was total of other cues in group and therefore not coded separately

### 7.3.2.2. Case by case analyses using frequency counts of verbal cues

Case by case analyses were conducted to assess the potential of each verbal cue for lie detection on the basis of the frequency counts. The cues considered in this analysis were those that had significantly discriminated between liars and truth-tellers. For each case (appeal), the score for each cue was compared with the overall cue mean. For the cue ‘positive emotion expressed towards the relative’ (a cue associated with honesty), frequency counts that fell above the cue mean for the honest appeals were classified as a ‘hit’, and frequency counts that fell below the cue mean for deceptive appeals were classified as a ‘hit’. For the remaining four cues (all associated with deception), frequency counts that fell above the cue mean for the deceptive appeals were classified as a ‘hit’, and frequency counts that fell below the

cue mean for the honest appeals were classified as a 'hit'. In this way, a percentage of 'hits' was calculated for each cue (i.e. cases combined). The cues had 'hit' rates between 78% and 100%, with a mean hit rate of 87%.

As in Study 2, the information from the cues was then combined to classify individual cases as deceptive or honest. For each case, the number of 'hits' on the cues was calculated (i.e. the number of times the appeal scored in the expected direction on each of the cues). Honest appeals had 'hit' rates between 80 - 100%, deceptive appeals had 'hit' rates between 80 - 100%, and the mean 'hit rate was 87%; see Table 7.3. As before, cases which scored above 50% 'hits' on the cues were categorised as correctly classified (the appeal scored in the expected direction on the majority of the cues), and cases that scored below 50% 'hits' on the cues were categorised as misclassified (the appeal did not score in the expected direction on the majority of the cues). In this way, all the appeals were correctly classified, supporting the hypothesis that using a combination of cues would result in high classification accuracy.

### ***7.3.3. Combined Case by Case Analyses Using Subjective Observer Ratings and Frequency Counts of Verbal Cues***

The next part of the case by case analyses assessed the potential of combining the results of Studies 2 and 3. The cues considered in this analysis were, therefore, those that had significantly discriminated between liars and truth-tellers in Studies 2 and 3. Case 'hit' rates using combined subjective and verbal cues ranged between 56% and 100%, with a mean 'hit' rate of 86%. As before, cases which scored above 50%

‘hits’ on the cues were categorised as correctly classified (the appeal scored in the expected direction on the majority of the cues), and cases that scored below 50% ‘hits’ on the cues were categorised as misclassified (the appeal did not score in the expected direction on the majority of the cues). In this way, all the appeals were correctly classified. See Table 7.3 for details.

Finally, an analysis was conducted to investigate whether any group of cues (subjective observer ratings, frequency counts of verbal cues, or subjective observer ratings and frequency counts of verbal cues combined) discriminated more successfully than any other group of cues between deceptive and honest appealers. To do this, the percentage ‘hit’ rates for each appealers were compared across the three groups of cues using a Friedman’s test. The result was not significant,  $\chi^2(2) = 4$ ,  $p=.135$ , indicating that there was no difference between subjective observer ratings of cues (median = 100), frequency counts of verbal cues (median = 80) and subjective observer ratings and frequency counts of verbal cues combined (median = 95), in successfully discriminating between deceptive and honest appealers.

Table 7.3: Case by case analyses: cues which discriminated significantly between honest and deceptive appeals: overall cue mean, score per case (appear), and overall 'hit' rate. For each case: 'hit' rate of subjective cues, 'hit' rate of verbal frequency count cues, and 'hit' rate of combined cues. All 'hit's related to overall cue means.

Cue	<i>M</i>	Honest					Deceptive				Cue hit rate
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	
<b><u>Subjective observer ratings</u></b>											
<b>Cues to Deception</b>											
Fake emotion	2.93	3.80	4.80	3.40	2.60	3.00	2.00	2.20	2.75	2.40	89%
Fake facial expression	2.99	3.80	4.40	3.00	3.00	3.00	2.60	2.20	2.75	2.60	100%
Performance	2.65	3.80	4.60	3.00	2.60	2.40	1.40	2.00	2.50	2.20	78%
High vocal pitch	3.34	4.20	3.60	4.20	4.00	3.60	1.40	3.40	3.00	3.20	89%
Lists	2.86	4.00	4.20	3.80	2.60	2.60	2.20	2.60	2.75	1.60	78%
Gaze aversion	2.64	3.80	4.40	3.80	2.00	3.00	1.40	2.00	2.50	1.60	89%
Lack of hope	3.10	3.60	N/A	3.60		N/A	2.80			2.40	100%
Lack of sympathy	2.90	4.20	4.80	3.00	2.80	2.80	2.00	N/A	N/A	2.40	78%
Creepy	3.11	4.20	4.80	3.20	3.00	3.60	2.20	2.40	2.25	3.00	89%
<b>Cues to honesty</b>											
Genuinely sad	2.90	3.80	4.40	3.40	3.00	2.40	2.00	2.20	3.00	2.60	78%
Facial expression	2.96	3.80	4.60	3.00	3.00	2.80	2.60	2.20	2.75	2.40	89%
Sad eyes	2.81	3.80	4.60	3.60	2.60	2.60	2.00	1.80	2.75	2.20	78%
Containing emotion	3.00	3.20	3.20	4.20	3.60	3.80	2.00	2.80	3.00 <sup>1</sup>	1.80	100%
Credible eye contact	2.75	3.80	4.40	3.60	2.40	2.40	1.80	2.00	2.75 <sup>1</sup>	2.20	75%
Hopeful	3.05	3.80	N/A	3.60	N/A	N/A	2.40	N/A	N/A	2.40	100%
Subjective observer ratings: case hit rate		100 %	100 %	100 %	38%	46%	100 %	92%	91%	100 %	



Cue	<i>M</i>	<u>Honest</u>					<u>Deceptive</u>				Cue hit rate
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	
<b><u>Verbal frequency counts</u></b>											
<b>Cues to Deception</b>											
Focus on others	1.42	<b>0.00</b>	<b>0.42</b>	<b>0.60</b>	<b>0.00</b>	<b>0.00</b>	1.33	0.66	<b>3.30</b>	<b>6.51</b>	78%
Deletion of 1 <sup>st</sup> person	0.40	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>1.77</b>	<b>0.66</b>	0.00	<b>1.18</b>	89%
Does not make sense	0.54	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.43</b>	<b>0.88</b>	<b>1.53</b>	<b>1.10</b>	<b>0.89</b>	100%
Repetition	1.17	<b>0.00</b>	1.26	1.20	<b>0.60</b>	<b>0.42</b>	<b>1.77</b>	<b>1.31</b>	<b>2.20</b>	<b>1.78</b>	78%
<b>Cue to honesty</b>											
Positive emotion for relative	1.96	0.00	<b>5.04</b>	<b>2.41</b>	<b>5.42</b>	<b>4.74</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	89%
Verbal freq. counts: case hit rate		80%	80%	80%	100 %	100%	80%	80%	80%	100 %	
Combined cues: case hit rate		95%	95%	95%	56%	61%	95%	89%	88%	100 %	

Case scores in bold are 'hits', in standard script are misses, <sup>1</sup> are unclassified (cue score is same as overall cue mean) and not included in hit rate. N/A Case in which body has been found so hope cues are not relevant.

#### **7.3.4. Verbal Frequency Counts: Summary**

The first aim of Study 3 was to investigate which of the verbal cues reported by the accurate participants in Study 1 (notwithstanding that some of these cues were also identified in previous research), reliably discriminated between honest and deceptive appealers. Of the 12 verbal cues suggested by Study 1, four were found to discriminate significantly between honest and deceptive appealers, with a further seven cues having large effect sizes in the expected direction, suggesting that the verbal cues identified by the accurate participants in Study 1 have some utility. Of

the verbal cues included which were suggested by previous research on high stakes deception, repetition was found to discriminate significantly between honest and deceptive appealers, and filled pauses had a large effect size in the expected direction. As hypothesised, the case by case analyses yielded high accuracy rates, suggesting that combining cues to investigate clusters of behaviour may be a productive approach in discriminating between liars and truth-tellers. Neither subjective observer ratings of cues alone, frequency counts of verbal cues alone, nor a combination of subjective observer ratings and frequency counts of verbal cues, proved to be more effective at discriminating between honest and deceptive appeals, suggesting that all channels of behaviour provide useful information.

#### **7.4. General Discussion**

To recap, the first aim of the present two studies was to investigate which of the cues reportedly used by participants accurate at detecting deception in Study 1, could reliably discriminate between honest and deceptive appeals. Overall, 19 individual cues suggested by the participants in Study 1 were found to discriminate between liars and truth-tellers, with a further 23 cues having large but non-significant effect sizes in the expected direction. This suggests that the methodology of investigating cues used by accurate participants may have some utility. Repetition, a cue suggested by previous research on high stakes lies, was also found to discriminate between deceptive and honest appeals.

The importance of assessing verbal behaviour with frequency counts as well as participant ratings became apparent as two cues (positive emotion expressed

towards the relative, and deletion of first person singular pronouns) which did not achieve inter-rater reliability in Study 2 (and were therefore dropped from further analysis), emerged as significant cues in Study 3. Two cues (focusing on others, and does not make sense) which had large but non-significant effect sizes in Study 2, were found to discriminate significantly between liars and truth-tellers in study 3, and a further two cues (using brutal language/detail, and avoiding using brutal language/detail) were found to have large effect sizes in the expected direction in the verbal frequency counts but not in the subjective observer ratings. It thus seems likely that some verbal cues are more suited to frequency counts than subjective observer ratings.

Of the cues investigated, those relating to emotional authenticity, distancing, and personal reaction emerged as particularly useful. In Study 2, of the ten cues related to emotional authenticity, seven discriminated significantly/marginally significantly between liars and truth-tellers, and a further two had large effect sizes in the expected direction. The salience of this category suggests two important implications: first, observers are able to distinguish between fake and genuine emotion, and second, fake emotion is a good indication of deception in this context. The deceptive appealers' use of fake emotion is presumably an attempt to appear honest by displaying the appropriate emotion of a concerned or grieving relative, and Ekman (2001) further suggests that liars may falsify a strong display of emotion as a strategy for masking a lie. The observers in Study 2 were able to make accurate overall judgements of fake and genuine emotion (on cues such as 'fake emotion', 'genuinely sad' and 'putting on a performance'), which is in line with previous findings that accuracy in judging liars is related to overall judgements of fake emotions (Vrij & Mann, 2001b). However, they were also able to make more

specific judgements on facial expression (fake facial expression, genuine facial expression, and sad eyes). As previously noted, this has relevance to the findings of ten Brinke and Porter's (2012) recent study in which the facial expressions of people making honest and deceptive appeals were analysed, frame by frame, using a facial coding system. They found that facial expressions related to disgust, distress, surprise and happiness discriminated between honest and deceptive appealers. The present results further suggest that untrained observers are able to detect differences between fake and genuine facial expressions in real time, and that these differences are useful in detecting deception in this context.

The finding that high vocal pitch was associated with deception is, as noted previously, in line with meta-analytic findings (DePaulo et al., 2003), and with previous research on high stakes lies (Hall, 1986), and is usually explained in terms of factors such as arousal, tension and anxiety (DePaulo et al., 2003; Vrij, 2004). However, it could be argued that anyone making an appeal for help with a missing or murdered relative, whether honest or deceptive, is likely to have elevated levels of tension and anxiety, and thus one might expect honest appealers to also have a high vocal pitch, at least to some degree. It could be that in the context of an honest appeal, sadness becomes the more salient emotional response, and sadness is typified by a low vocal pitch (Murray & Arnott, 1993).

Cues related to distancing were important indicators of deception in the context of appeals, and, as previously discussed, this is in line with the theoretical suggestions of several researchers (Buller & Burgoon, 1996; Ekman, 2001; Zuckerman et al., 1981) that liars will communicate in more evasive and indirect ways than truth-tellers, in an attempt to distance themselves from their own deceptive communication.

In this context, the emergence of both gaze aversion and credible eye contact as significantly/marginally significantly discriminating variables between liars and truth-tellers, with large effect sizes, is particularly interesting, given that previous research in low stakes deception situations has shown that gaze aversion is not related to deception (DePaulo et al., 2003), or even that liars may increase eye contact in an attempt to appear honest, or monitor how their receiver is responding to their lies (Mann et al., 2012). However, as noted earlier, there are a variety of a priori reasons why we might expect gaze aversion to increase with deception in high stakes contexts; these include reactions to feelings of shame and guilt (DePaulo et al., 2003; Ekman, 2001), emotions that may be particularly relevant in the context of somebody who is lying about killing their own relative. However, particularly in the present context, it is possible that increased gaze aversion may also be used as a form of distancing behaviour. Honest appeals are likely to be a genuine attempt to communicate with the public and ask for help, hence honest appealers are more likely to behave in a way that facilitates this communication; for example, by using direct eye contact. Nevertheless, the motivation of deceptive appealers is clearly not the same, as they do not actually want the public to assist in finding out what happened to their relative. Deceptive appealers may, therefore, produce behaviours that inhibit direct communication, such as, gaze aversion.

Two of the cues that could be construed as related to distancing that discriminated between liars and truth-tellers may have specific relevance to the particular context of making an appeal; these are, focusing on others, and lists, both of which may serve to distance the liar from the subject of their speech (i.e. their relative). The non-significant but large effect sizes for the cues ‘equivocation’, and ‘irrelevant/generic statements’, lend some further support to the theoretical proposal

that deception may be associated with linguistic constructions that distance the liar from the subject of their speech. Conversely, the cue ‘positive emotion expressed towards the relative’, which was associated with honesty in the specific context of appeals, could be regarded as functioning in the opposite way, so that rather than dissociating the appealer from their relative, it indicates engagement.

It has been suggested by Sapir (1987), as part of his scientific content analysis technique, that the deletion of first person singular pronouns is also indicative of a lack of commitment and responsibility for what is being said. However, findings regarding this cue are mixed; as noted in Chapter 6, some studies have supported this proposal (for example, Newman et al., 2003; Dilmon, 2009), whilst others have not (for example, Porter & Yuille, 1996); also, meta-analytic findings do not offer support for the idea (DePaulo et al., 2003). It may be that the deletion of first person singular pronouns is a cue that only emerges in specific contexts.

Cues related to the personal reaction of the observers produced results which are notable for two reasons. First, even though the cues were entirely subjective, seven of the eight cues related to personal reaction in Study 2 achieved significant inter-rater reliability, suggesting that acceptable levels of consensus may exist even for highly subjective cues. Second, all the cues analysed had large effect sizes in the expected direction, suggesting that personal reactions may be useful in discriminating between liars and truth-tellers.

As noted in Chapter 6, there are several possible explanations for this effect. For example, Hatfield et al.’s (1992) theory of emotional contagion suggests that our tendency to automatically mimic the facial, vocal and postural behaviours of people we are watching results in feedback, which causes us to experience pale reflections

of the emotions felt by the people we are watching. Moreover, recent research on mirror neurons suggests that we have mirroring mechanisms with which we respond to other people's experiences (Mukamel, Ekstrom, Kaplan, Iacoboni & Fried, 2010). The present results suggest that this emotional synchronisation effect may occur for genuine emotions but not for faked emotions: observers demonstrated a significant lack of sympathy for deceptive appealers, but not honest appealers. Deceptive appealers were also rated as being significantly more 'creepy' than honest appealers. This could be construed as in line with Zuckerman et al.'s (1981) proposal that, because they express negative feelings of guilt or fear, liars may appear and sound more unpleasant than truth-tellers. On similar lines, De Paulo et al.'s (2003) self-presentation theory predicts that due to moral qualms, anxiety and negativity, a lack of personal investment, and attempts at behavioural control, liars will appear less pleasant than truth-tellers. Accordingly, in their meta-analysis they found that all cues assessing pleasantness produced results in the predicted direction. It would seem, therefore, that subjective personal reactions may have some value in discriminating between liars and truth-tellers.

Regarding cues relating to a lack of emotion, it is interesting to note that whilst observers were good at distinguishing fake emotion from genuine emotion, a perceived lack of emotion proved to be less predictive overall. Previous researchers have suggested that a perceived lack of emotion may result from liars' efforts to control their behaviour and prevent leakage, and this attempt at impression management may result in their behaviour appearing rigid or inhibited (Vrij, 2004). However, in the present research, the cue 'containing emotion', which involves controlling behaviour, was related, as expected, to honesty, not to deception. This finding does not sit comfortably with the theory that liars control their behaviour

more than truth-tellers. It is possible that, in the specific context of making appeals, attempts to control behaviour may be the result of factors other than deception (for example, being uncomfortable on camera, or, in the case of honest appealers, a determination to get their plea for help across). On the whole, it would seem that in the specific context of appeals, fake emotion may be more important in the detection of deception than signs of controlling behaviour and lack of emotion.

Cues relating to plausibility also emerged as useful in discriminating between deceptive and honest appeals. The finding that the cue 'does not make sense' discriminated between deceptive and honest appeals is in line with both meta-analytic findings that lies make less sense than truths (DePaulo et al, 2003), and with the more specific finding that in 911 homicide calls, conflicting facts predicted deception (Harpster et al., 2009). Furthermore, a particular behaviour emerged in Study 2 that may have specific relevance to the context of making an appeal. In cases where the relative is missing and no body had been found, honest appealers were rated significantly higher than deceptive appealers on the cue 'hopeful', and conversely deceptive appealers were rated significantly higher than honest appealers on the cue 'lack of hope'. While this cue is clearly context-specific, it may be applicable in other specified contexts; for example, Harpster et al. (2009) found in their analysis of 911 homicide calls, that acceptance of death was significantly related to guilt, and this is very similar to a lack of hope of finding the relative alive in an appeal.

The results of the case by case analyses were also promising, suggesting that, as hypothesised, the more cues to deception that a person exhibits, the more likely he/she is to be lying. In deception research in general, few cues are found which consistently and reliably discriminate between liars and truth-tellers (DePaulo et al,



2003), presumably because individual liars are unlikely to produce all the behaviours associated with deception and thus significant findings tend to be patchy. However, the examination of cues in combination to investigate clusters of behaviour may be more successful in discriminating between liars and truth-tellers, as this approach allows for individual variations in behaviour, whilst still detecting a cumulative effect. For the assessment of whether an individual case was honest or deceptive, a combination of the subjective observer ratings and frequency counts of verbal cues produced high 'hit' rates, with seven of the nine cases yielding 'hit' rates above 87%. Interestingly, the two cases (cases 4 and 5) which yielded low 'hit' rates in Study 2 (38% and 46% respectively), both yielded 'hit' rates of 100% in Study 3. This supports the proposal of a number of previous researchers (for example, O'Sullivan & Ekman, 2004), that it is important to utilise multiple channels when attempting to detect deception.

As hypothesised, the observers' credibility judgements in Study 2 were accurate well above the level of chance, and the average accuracy of 75% is in line with other studies using real life, high stakes lies as stimulus materials (Mann & Vrij, 2006; Mann et al., 2006; Vrij et al., 2006). This suggests that the use of real life, high stakes stimulus materials may have an effect on observers' ability to discriminate between honesty and deception. However, the possibility that the observers achieved higher than usual accuracy rates because they were paying attention to cues that may be helpful in discriminating between deceptive and honest appeals, cannot be ruled out. Furthermore, the sample size of observers was clearly too small to draw any firm conclusions regarding accuracy. These issues are addressed more fully in Study 7 of the present thesis.

It can be noted also that, of the cues found to discriminate significantly between liars and truth-tellers, several were previously uninvestigated cues that may have relevance to the specific context of making an appeal: putting on a performance, lists, a lack of hope, creepy, a lack of sympathy (all associated with deception); and containing emotion, hopeful and positive emotion expressed towards the relative (all related to honesty). Of the non-significant cues that nevertheless had large effect sizes, there were also several that were new and may relate specifically to appeals: pretending to cry, cold, focusing on others, defensive, and using brutal language (all related to deception); and calmness appropriate to speech content, normal, feeling sorry for the appellant, feeling the appellants' pain, and avoiding brutal language, (all related to honesty).

However, clearly, a limitation of the present studies was the small sample size of appeals. The small sample size allowed the possibility that one or two cases (either deceptive or honest) may have unduly influenced the findings, and also resulted in a lack of power, so that cues with large effect sizes yielded non-significant results. Despite this, 20 cues were found to discriminate significantly between liars and truth-tellers, and a further 23 had large effect sizes in the expected direction, suggesting that in the context of making appeals, discernible patterns of deceptive and honest behaviour may exist. Furthermore, the high accuracy rates yielded in the case by case analyses suggest that combining these cues may be useful for classifying individual cases as deceptive or honest. Consequently, it was decided to repeat the investigation on a different sample of observers and with a larger sample of appeals.

# **Chapter 8**

## **Study 4: Subjective Observer Ratings of Cues to Deception in Appeals**

### **8.1 Introduction**

Given the issues concerning the robustness of the findings in Study 2, the first aim of the study described in this chapter was to replicate and extend the trends identified in Study 2 on a new and larger sample set, and using a new panel of observers. The second aim of the study was to investigate which of some other subjective cues to deception, identified in previous research in other high stakes contexts, would also have utility in the context of appeals. The third aim was, again, to utilise case by case analyses to assess the potential of each cue for deception detection, and also the potential for the combination of cues to classify individual cases as honest or deceptive. Finally, as in Study 2, it was hypothesised that the use of high stakes stimulus material would result in accuracy levels above chance, and elimination of truth bias.

## **8.2 Method**

### **8.2.1. *Participants***

Five adult observers (one male and four female) were used to rate the appeals. Ages of the observers ranged from 18-21, and all were undergraduate psychology students participating for course credit and recruited using the University online participation system. None had previously received any formal training in lie detection techniques.

### **8.2.2. *Materials and Procedure***

Participants attended as a group at a psychology laboratory at the University of Liverpool, and the session lasted approximately two hours. The procedure was the same as for Study 2, except observers were shown 32 appeals; 16 honest and 16 deceptive appeals. This was a larger sample set than is common in high stakes deception studies; some studies have used only one ‘sender’ (for example, Villar et al., 2011; Vrij & Mann, 2001a), and a body of research has been based on police interviews of 16 suspects (Mann et al., 2002; Mann & Vrij, 2006; Mann et al., 2006). The criteria for the selection of cases included in the study were described in Chapter 5. See Appendix 3 for case summaries.

As before, the response sheet given to each observer contained one statement for each cue. There were 28 statements, one for each predetermined cue, and statements indicated cues to both deception and to honesty (for example, ‘This

person seems creepy’, and ‘This person is genuinely sad’). Observers were asked to indicate their agreement with each statement on a scale of 1 (*strongly disagree*) to 5 (*strongly agree*). At the end of each response sheet, observers were asked to indicate whether they thought the appealer was deceptive or honest. Cues included on the response sheet were nonverbal cues identified in Study 2 of the present thesis as discriminating significantly between deceptive and honest appeals, or having large effect sizes in the expected direction (verbal cues were investigated separately in Study 5). It can be noted, therefore, that actuarial rather than theoretical prediction provided the primary rationale for the inclusion of these cues at this stage; possible theoretical explanations for their efficacy are explored later. Due to the lack of previous research in this area, and also for reasons discussed in Chapter 7, it was decided to include non-significant cues with large effect sizes in the expected direction, to minimise the risk of Type II errors occurring.

Also five further subjective cues identified in previous research in high stakes deception were included. The cue ‘urgency’ was included due to the finding by Harpster et al. (2009), in their study of 911 homicide calls, that urgency of the plea for help was related to honesty, and it was considered that this behaviour may have direct relevance to the context of appeals. In the same study, vocal modulation, including the degree to which the voice sounds personal and expressive, was also related to honesty; as meta-analytic findings (DePaulo et al., 2003) also indicate that the cue ‘personal and expressive voice’ is related to honesty, it was included as a cue in the present study. On similar lines, liars have been found to sound less certain and direct across high stake situations (Koper & Sahlman, 1991; Wright Whelan, 2009), and this cue also emerged as significant in a meta-analysis (DePaulo et al., 2003), and it was therefore decided to include it in the present study. Speech hesitation has

been found to relate to deception in previous research on high stakes deception (Wright Whelan, 2009), as has pausing (Mann et al., 2002; Vrij & Mann, 2001a), and it was therefore included in the present study. Finally, general plausibility has been related to honesty in high stakes situations (Koper & Sahlman, 1991), and also meta-analytically (DePaulo et al., 2003), and was included here. The cue ‘voice quivering with emotion’ was noticed by the researcher whilst coding data for Study 5, and although, to the author’s knowledge, no previous research has identified this cue specifically as relating to veracity, it relates to Harpster et al.’s (2009) vocal modulation cue.

The cue genuine facial expression, which in Study 2 was related to honesty, was dropped as it was considered to be the negative equivalent, and, therefore, repetition of, fake facial expression. Similarly, the cue genuine behaviours was dropped as it was considered to be the negative equivalent, and, therefore, repetition, of, behaving unnaturally. Verbal cues that had emerged in Study 2 as having potential utility in discriminating between deception and honesty were not included in the present study, as they were analysed separately using frequency counts on transcripts of the appeals (see Chapter 9). See Appendix 7 for the full response sheet.

## **8.3 Results**

### **8.3.1. *Inter-rater Reliability***

To assess inter-rater reliability between the five observers, Kendall’s W was calculated for each cue. Inter-rater reliability was not significant for seven of the cues

( $p > .05$ ), which were dropped from further analysis (pretending to cry, no emotional variation, gaze aversion, defensive, containing emotion, calmness appropriate to speech content, and credible eye contact). One cue, (no sympathy for the appealer) had marginally significant inter-rater reliability ( $p = .054$ ), and was retained for further analysis; see Table 8.1.

### **8.3.2. *Analyses of Cues***

The mean of the five observers' scores for each cue in each appeal was calculated, so that each appeal had a single score for each cue, and these scores were used in all further analyses. Two of the observers were familiar with the outcome of one of the 32 appeals, and as ratings of the cues were required to be independent of the observers' knowledge of the veracity of the appeals, ratings of the cues for this one appeal were calculated from the scores of the other three observers who were not familiar with the person in the appeal.

A MANOVA with Veracity (Honest/Deceptive) as the between-subjects independent variable, was conducted on the mean cue scores for the cues expected to be related to deception. Means, standard deviations,  $F$  ratios,  $p$  values and effect sizes are presented in Table 8.1. Using Pillai's Trace, the multivariate test yielded a significant result,  $F(1,21) = 2.49$ ,  $p = .038$ . Follow-up univariate analyses showed significant effects for the following cues, all with means in the expected direction: fake emotion, fake facial expression, putting on a performance, dislike for the appealer, creepy, behaving unnaturally, and no sympathy for the appealer.

Levene's test for homogeneity of variance was significant for the cue fake emotion, and although the MANOVA is fairly robust to violations of the standard assumptions with fixed levels of the independent variable and fairly equal sample sizes (Glass & Stanley, 1970), Welch's Test was run on this cue as a check; it produced a result comparable to the univariate test ( $p = .001$ ).

A similar MANOVA with Veracity as a between-subjects independent variable, was also conducted on the scores for the cues expected to be related to honesty. Means, standard deviations,  $F$  ratios,  $p$  values and effect sizes are presented in Table 8.1. Using Pillai's Trace, the multivariate test showed a result that approached significance,  $F(1, 19) = 2.10$ ,  $p = .072$ ; given the result approached significance, and a preponderance of large effect sizes, follow-up univariate analyses were conducted, nonetheless. These showed significant effects for the following cues, all with means in the expected direction: genuinely sad, sad eyes, genuine and heartfelt appeal, feel the appealer's pain, feel sorry for the appealer, normal, urgency, personal and expressive voice, plausible, and voice quivering with emotion.

Levene's test for homogeneity of variance was significant for the cues sad eyes, feel sorry for the appealer, and plausible. However, Welch's Test produced results comparable to the univariate tests; sad eyes  $p = .015$ , feel sorry for the appealer  $p = .002$ , plausible  $p = .005$ .

In total, therefore, 17 cues discriminated successfully between deceptive and honest appealers, and, as Table 8.1 shows, all had large effect sizes in this respect.



Table 8.1: *Kendall's Ws for cues rated subjectively by observers; means, SDs, CIs, F ratios and p values, and effect sizes for cues which achieved significant inter-rater reliability*

Cue	Kendall's W ( <i>p</i> )	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
<b><u>Cues to deception</u></b>					
Fake emotion*	.55 ( <i>p</i> = .014)	<b>3.33</b> (0.79) 3.00 – 3.67	<b>2.49</b> (0.48) 2.15 – 2.83	<i>F</i> (1,31) = 13.50 <i>p</i> = .001	<i>d</i> = 1.29
Fake facial expression	.58 ( <i>p</i> = .007)	<b>3.35</b> (0.73) 3.00 – 3.69	<b>2.49</b> (0.60) 2.14 – 2.83	<i>F</i> (1,31) = 13.25 <i>p</i> = .001	<i>d</i> = 1.29
No sympathy	.48 ( <i>p</i> = .054)	<b>3.24</b> (0.65) 2.96 – 3.51	<b>2.56</b> (0.41) 2.30 – 2.85	<i>F</i> (1,31) = 11.96 <i>p</i> = .002	<i>d</i> = 1.25
Performance	.50 ( <i>p</i> = .039)	<b>3.38</b> (0.71) 3.05 – 3.70	<b>2.61</b> ( 0.52) 2.29 – 2.94	<i>F</i> (1,31) = 11.92 <i>p</i> = .002	<i>d</i> = 1.24
Dislike	.54 ( <i>p</i> = .016)	<b>3.42</b> (0.65) 3.13 – 3.71	<b>2.74</b> (0.46) 2.45 – 3.02	<i>F</i> (1,31) = 11.87 <i>p</i> = .002	<i>d</i> = 1.21
Creepy	.53 ( <i>p</i> = .019)	<b>3.30</b> (0.71) 2.99 – 3.60	<b>2.60</b> (0.47) 2.29 – 2.91	<i>F</i> (1,31) = 10.61 <i>p</i> = .003	<i>d</i> = 1.16
Behaving unnaturally	.50 ( <i>p</i> = .035)	<b>3.40</b> (0.70) 3.06 – 3.74	<b>2.76</b> (0.63) 2.42 – 3.10	<i>F</i> (1,31) = 7.38 <i>p</i> = .011	<i>d</i> = 0.96
Cold	.60 ( <i>p</i> = .004)	3.25 (0.76) 2.84 – 3.66	2.73 ( 0.82) 2.31 – 3.14	<i>F</i> (1,31) = 3.50 <i>p</i> = .071	<i>d</i> = 0.66
High vocal pitch	.67 ( <i>p</i> = .001)	3.02 (0.72) 2.63 – 3.41	3.23 (0.82) 2.83 – 3.62	<i>F</i> (1,31) = 0.56 <i>p</i> = .459	<i>d</i> = 0.27
Pretending to cry	.34 ( <i>p</i> = .448)				
No emotional variation	.46 ( <i>p</i> = .084)				
Gaze aversion	.43 ( <i>p</i> = .124)				
Defensive	.63 ( <i>p</i> = .155)				
Hesitant	.55 ( <i>p</i> = .014)	2.68 (0.72) 2.34 – 3.02	2.85 (0.60) 2.51 – 3.19	<i>F</i> (1,31) = 0.51 <i>p</i> = .483	

Cue	Kendall's W ( <i>p</i> )	<u>Deceptive</u> Mean (SD) CI	<u>Honest</u> Mean (SD) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
<b><u>Cues to honesty</u></b>					
Feel sorry for the appealer*	.57 ( <i>p</i> = .009)	<b>2.68</b> ( 0.78) 2.35 – 3.01	<b>3.48</b> (0.51) 3.14 – 3.81	<i>F</i> (1,31) = 11.80 <i>p</i> = .002	<i>d</i> = 1.21
Genuinely sad	.61 ( <i>p</i> = .003)	<b>2.65</b> (0.85) 2.29 – 3.02	<b>3.51</b> (0.57) 3.15 – 3.88	<i>F</i> (1,31) = 11.35 <i>p</i> = .002	<i>d</i> = 1.19
Voice quivering with emotion	.56 ( <i>p</i> = .011)	<b>2.50</b> (0.81) 2.11 – 2.90	<b>3.41</b> (0.74) 3.02 – 3.81	<i>F</i> (1,31) = 11.00 <i>p</i> = .002	<i>d</i> = 1.17
Genuine/heartfelt appeal	.66 ( <i>p</i> = .001)	<b>2.55</b> (0.87) 2.17 – 2.94	<b>3.40</b> (0.63) 3.01 – 3.79	<i>F</i> (1,31) = 9.94 <i>p</i> = .004	<i>d</i> = 1.12
Normal	.49 ( <i>p</i> = .044)	<b>2.74</b> (0.64) 2.46 – 3.02	<b>3.35</b> (0.44) 3.07 – 3.63	<i>F</i> (1,31) = 10.03 <i>p</i> = .004	<i>d</i> = 1.11
Plausible*	.51 ( <i>p</i> = .032)	<b>2.71</b> (0.68) 2.44 – 2.99	<b>3.30</b> (0.36) 3.02 – 3.58	<i>F</i> (1,31) = 9.42 <i>p</i> = .005	<i>d</i> = 1.08
Personal & expressive voice	.49 ( <i>p</i> = .043)	<b>2.76</b> (0.65) 2.43 – 3.09	<b>3.40</b> (0.63) 3.07 – 3.73	<i>F</i> (1,31) = 7.86 <i>p</i> = .009	<i>d</i> = 1.00
Feel the appealer's pain	.56 ( <i>p</i> = .010)	<b>2.70</b> (0.75) 2.36 – 3.05	<b>3.35</b> (0.58) 3.01 – 3.69	<i>F</i> (1,31) = 7.37 <i>p</i> = .011	<i>d</i> = 0.97
Sad eyes*	.60 ( <i>p</i> = .004)	<b>2.55</b> (1.05) 2.12 – 2.98	<b>3.33</b> (0.53) 2.90 – 3.75	<i>F</i> (1,31) = 6.89 <i>p</i> = .013	<i>d</i> = 0.94
Urgency	.57 ( <i>p</i> = .009)	<b>3.00</b> (0.83) 2.59 – 3.41	<b>3.65</b> ( 0.77) 3.24 – 4.06	<i>F</i> (1,31) = 5.27 <i>p</i> = .029	<i>d</i> = 0.81
Containing emotion	.26 ( <i>p</i> = .797)				
Calm appropriate to speech content	.47 ( <i>p</i> = .061)				
Credible eye contact	.42 ( <i>p</i> = .158)				
Certain & direct	.51 ( <i>p</i> = .033)	3.28 (0.70) 2.99 – 3.58	3.45 (0.42) 3.16 – 3.74	<i>F</i> (1,31) = 0.68 <i>p</i> = .417	

Means in bold differ significantly/marginally significantly from each other, \*Levene's significant

### **8.3.3. Case by Case Analyses**

As in Studies 2 and 3, and in accordance with previous studies (Ekman et al., 1991; Leal et al., 2010), case by case analyses were conducted to assess the potential of each cue for lie detection on the basis of the subjective observers' ratings. The cues considered in these analyses were those that had significantly discriminated between liars and truth-tellers. As before, for each case, the score for each cue was compared with the overall cue mean. For deceptive appeals, cue scores that fell above the cue mean on cues to deception, and below the cue mean on cues to honesty, were classified as a 'hit'. Similarly, for honest appeals, cue scores that fell above the cue mean on cues to honesty, and below the cue mean on cues to deception, were classified as a 'hit'. In this way, a percentage of 'hits' was calculated for each cue (i.e. cases combined). The cues had 'hit' rates between 75% (creepy, no sympathy for the appealer, and voice quivering with emotion), and 63% (unnatural, and feel the appealer's pain). See Table 8.2.

The percentage of 'hits' was also calculated for each case (i.e. cues combined). Honest appeals had 'hit' rates between 0 - 100%, and deceptive appeals had 'hit' rates between 24 - 100% (see Table 8.2). As before, cases which scored above 50% 'hits' on the cues were categorised as correctly classified (the appeal scored in the expected direction on the majority of the cues), and cases that scored below 50% 'hits' on the cues were categorised as misclassified (the appeal did not score in the expected direction on the majority of the cues). In this way, 10 of the 16 deceptive appeals were correctly classified, and 12 of the 16 honest appeals were correctly classified. Overall, 69% of appeals were correctly classified.

Table 8.2: Case by case analyses; cues rated by observers which discriminated significantly between truthful and deceptive appeals. Overall cue mean, score per case (appealer), and overall 'hit' rate. For each case: 'hit' rate of cues. All 'hit's related to overall cue means.

	Fake emoti on  ( <i>M</i> = 2.91)	Fake facial expres sion  ( <i>M</i> = 2.92)	Perfor mance  ( <i>M</i> = 2.99)	Unnat ural  ( <i>M</i> = 3.08)	creep y  ( <i>M</i> = 2.95)	Dislik e  ( <i>M</i> = 3.08)	No sympa thy  ( <i>M</i> = 2.91)	Sad eyes  ( <i>M</i> = 2.94)	Genui nely sad  ( <i>M</i> = 3.08)	Genui ne / heartf elt  ( <i>M</i> = 2.98)	Urgen cy  ( <i>M</i> = 3.33)	Plausi ble  ( <i>M</i> = 3.01)	Norm al  ( <i>M</i> = 3.04)	Perso nal & expres sive  ( <i>M</i> = 3.08)	Vocal quiver  ( <i>M</i> = 2.96)	Feel pain  ( <i>M</i> = 3.03)	Feel sorry  ( <i>M</i> = 3.08)	Case 'hit' rate
<b>Deceptive appeals</b>																		
Case 10	<b>4.00</b>	<b>3.60</b>	<b>3.80</b>	<b>3.60</b>	2.60	<b>3.40</b>	<b>3.60</b>	<b>2.20</b>	<b>2.00</b>	<b>1.80</b>	<b>2.20</b>	<b>2.60</b>	<b>2.60</b>	<b>2.20</b>	<b>1.40</b>	<b>2.40</b>	<b>2.20</b>	94%
Case 11	<b>4.20</b>	<b>4.00</b>	<b>4.20</b>	<b>4.40</b>	<b>4.40</b>	<b>4.00</b>	<b>4.00</b>	<b>1.20</b>	<b>1.40</b>	<b>1.20</b>	<b>3.20</b>	<b>1.60</b>	<b>2.20</b>	<b>1.80</b>	<b>1.40</b>	<b>1.60</b>	<b>2.20</b>	100%
Case 12	<b>4.40</b>	<b>4.40</b>	<b>4.60</b>	<b>4.80</b>	<b>3.80</b>	<b>4.20</b>	<b>4.00</b>	<b>1.40</b>	<b>1.40</b>	<b>1.40</b>	<b>1.60</b>	<b>2.00</b>	<b>2.00</b>	<b>2.40</b>	<b>1.40</b>	<b>1.60</b>	<b>1.40</b>	100%
Case 13	2.80	2.80	<b>3.00</b>	2.60	<b>3.20</b>	3.00	2.80	3.00	3.40	3.60	<b>3.00</b>	3.40	<b>3.00</b>	<b>3.00</b>	<b>2.60</b>	3.20	3.60	35%
Case 14	2.80	2.80	2.60	<b>3.60</b>	<b>3.00</b>	<b>3.20</b>	<b>3.00</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	<b>2.80</b>	3.40	3.60	<b>2.60</b>	<b>2.40</b>	<b>2.80</b>	<b>2.80</b>	71%
Case 15	<b>3.20</b>	<b>3.00</b>	<b>3.20</b>	<b>3.40</b>	<b>3.60</b>	<b>3.20</b>	<b>3.20</b>	3.00	<b>2.80</b>	<b>2.60</b>	<b>2.60</b>	<b>2.00</b>	<b>2.40</b>	<b>2.40</b>	<b>2.60</b>	<b>2.80</b>	<b>3.00</b>	94%
Case 16	2.60	2.60	2.80	3.00	<b>4.20</b>	3.00	2.60	4.00	3.60	3.40	4.40	3.20	<b>1.80</b>	3.60	3.60	3.40	3.60	12%
Case 17	<b>4.33</b>	<b>4.33</b>	<b>4.00</b>	<b>4.00</b>	<b>3.33</b>	<b>4.33</b>	<b>4.00</b>	<b>1.00</b>	<b>1.67</b>	<b>1.67</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>1.67</b>	<b>1.67</b>	<b>1.67</b>	100%
Case 18	<b>3.60</b>	<b>3.80</b>	<b>3.80</b>	<b>3.20</b>	<b>3.20</b>	<b>3.40</b>	<b>3.60</b>	<b>1.40</b>	<b>2.60</b>	<b>2.60</b>	3.40	<b>2.20</b>	<b>2.80</b>	<b>2.80</b>	<b>2.60</b>	<b>2.40</b>	<b>2.60</b>	94%
Case 19	2.80	<b>3.20</b>	<b>3.00</b>	3.00	2.20	<b>3.20</b>	<b>3.00</b>	3.40	3.20	3.00	<b>2.60</b>	3.20	3.20	3.20	3.40	3.20	3.20	29%
Case 21	2.60	2.60	2.40	3.00	2.60	2.80	2.60	3.60	3.40	3.60	3.40	3.40	3.80	3.40	3.00	3.60	3.40	0%
Case 20	2.60	2.40	2.80	2.80	2.80	2.60	2.60	3.60	3.60	3.40	4.20	3.60	3.60	3.40	3.20	3.60	3.40	0%

	Fake emoti on  ( <i>M</i> = 2.91)	Fake facial expres sion  ( <i>M</i> = 2.92)	Perfor mance  ( <i>M</i> = 2.99)	Unnat ural  ( <i>M</i> = 3.08)	creep y  ( <i>M</i> = 2.95)	Dislik e  ( <i>M</i> = 3.08)	No sympa thy  ( <i>M</i> = 2.91)	Sad eyes  ( <i>M</i> = 2.94)	Genui nely sad  ( <i>M</i> = 3.08)	Genui ne / heartf elt  ( <i>M</i> = 2.98)	Urgen cy  ( <i>M</i> = 3.33)	Plausi ble  ( <i>M</i> = 3.01)	Norm al  ( <i>M</i> = 3.04)	Perso nal & expres sive  ( <i>M</i> = 3.08)	Vocal quiver  ( <i>M</i> = 2.96)	Feel pain  ( <i>M</i> = 3.03)	Feel sorry  ( <i>M</i> = 3.08)	Case 'hit' rate
Case 22	1.80	2.20	2.20	2.00	2.20	2.20	1.80	4.40	4.20	4.00	4.40	3.60	3.40	4.20	4.00	4.00	3.80	0%
Case 23	<b>4.20</b>	<b>4.20</b>	<b>4.00</b>	<b>4.00</b>	<b>3.80</b>	<b>4.40</b>	<b>3.60</b>	<b>1.80</b>	<b>2.20</b>	<b>1.80</b>	<b>3.20</b>	<b>2.20</b>	<b>2.80</b>	<b>2.40</b>	<b>1.80</b>	<b>2.20</b>	<b>2.20</b>	100%
Case 24	<b>3.60</b>	<b>3.80</b>	<b>3.80</b>	<b>3.60</b>	<b>3.40</b>	<b>3.80</b>	<b>3.40</b>	<b>2.20</b>	<b>2.20</b>	<b>2.00</b>	<b>2.60</b>	<b>2.60</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	100%
Case 25	<b>3.80</b>	<b>3.80</b>	<b>3.80</b>	<b>3.40</b>	<b>4.40</b>	<b>4.00</b>	<b>4.00</b>	<b>2.20</b>	<b>2.20</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	<b>2.20</b>	<b>2.20</b>	<b>2.60</b>	<b>2.20</b>	<b>1.80</b>	100%
Cue 'hit' rate; deceptive appeals	56%	63%	69%	63%	69%	69%	69%	56%	63%	63%	69%	56%	69%	69%	69%	63%	63%	
<b><u>Honest appeals</u></b>																		
Case 26	3.00	3.40	3.00	3.20	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	2.60	3.00	<b>3.00</b>	2.60	<b>3.20</b>	<b>3.60</b>	<b>3.20</b>	2.20	<b>3.20</b>	<b>3.40</b>	53%
Case 27	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.60</b>	<b>2.80</b>	<b>2.60</b>	<b>3.20</b>	2.80	<b>3.00</b>	2.60	<b>3.20</b>	<b>3.60</b>	2.80	2.60	2.80	3.00	65%
Case 28	<b>2.60</b>	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.60</b>	<b>2.60</b>	<b>2.60</b>	2.60	<b>3.40</b>	<b>3.00</b>	<b>3.40</b>	<b>3.60</b>	<b>3.60</b>	<b>3.80</b>	<b>3.20</b>	3.00	<b>3.40</b>	88%
Case 29	<b>1.80</b>	<b>1.40</b>	<b>2.20</b>	<b>2.20</b>	<b>2.20</b>	<b>2.40</b>	<b>2.00</b>	<b>4.20</b>	<b>4.20</b>	<b>4.40</b>	<b>4.60</b>	<b>3.40</b>	<b>3.60</b>	<b>4.40</b>	<b>4.40</b>	<b>3.80</b>	<b>4.20</b>	100%
Case 30	<b>2.80</b>	<b>2.60</b>	<b>2.60</b>	3.40	3.00	<b>2.60</b>	<b>2.80</b>	<b>3.40</b>	<b>3.40</b>	<b>3.40</b>	<b>3.40</b>	<b>3.20</b>	<b>3.20</b>	<b>3.20</b>	<b>3.80</b>	<b>3.40</b>	<b>3.80</b>	88%
Case 31	3.00	3.00	3.20	3.20	<b>2.40</b>	<b>3.00</b>	3.00	<b>3.00</b>	2.60	2.60	3.20	3.00	<b>3.40</b>	3.00	2.40	2.80	3.00	24%

	Fake emoti on  ( <i>M</i> = 2.91)	Fake facial expres sion  ( <i>M</i> = 2.92)	Perfor mance  ( <i>M</i> = 2.99)	Unnat ural  ( <i>M</i> = 3.08)	creep y  ( <i>M</i> = 2.95)	Dislik e  ( <i>M</i> = 3.08)	No sympa thy  ( <i>M</i> = 2.91)	Sad eyes  ( <i>M</i> = 2.94)	Genui nely sad  ( <i>M</i> = 3.08)	Genui ne / heartf elt  ( <i>M</i> = 2.98)	Urgen cy  ( <i>M</i> = 3.33)	Plausi ble  ( <i>M</i> = 3.01)	Norm al  ( <i>M</i> = 3.04)	Perso nal & expres sive  ( <i>M</i> = 3.08)	Vocal quiver  ( <i>M</i> = 2.96)	Feel pain  ( <i>M</i> = 3.03)	Feel sorry  ( <i>M</i> = 3.08)	Case 'hit' rate
Case 32	<b>1.80</b>	<b>1.80</b>	<b>2.00</b>	<b>2.40</b>	<b>2.20</b>	<b>1.80</b>	<b>2.00</b>	<b>4.00</b>	<b>4.40</b>	<b>4.40</b>	<b>4.80</b>	<b>4.00</b>	<b>3.80</b>	<b>4.00</b>	<b>4.00</b>	<b>4.20</b>	<b>4.00</b>	100%
Case 33	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>2.40</b>	<b>2.20</b>	<b>2.00</b>	<b>3.20</b>	<b>3.80</b>	<b>3.80</b>	<b>3.40</b>	<b>3.80</b>	<b>3.80</b>	<b>3.80</b>	<b>4.00</b>	<b>4.00</b>	<b>3.80</b>	100%
Case 34	3.00	3.00	3.00	3.40	3.40	3.40	<b>2.80</b>	<b>3.00</b>	<b>3.20</b>	<b>3.00</b>	<b>4.60</b>	2.80	2.40	3.00	<b>3.40</b>	<b>3.20</b>	3.00	41%
Case 35	<b>2.40</b>	<b>2.40</b>	<b>2.40</b>	3.80	<b>2.80</b>	3.40	<b>2.80</b>	<b>3.20</b>	3.00	2.60	3.00	<b>3.40</b>	3.00	2.40	2.60	2.80	<b>3.20</b>	47%
Case 36	<b>2.80</b>	<b>2.80</b>	3.40	<b>2.80</b>	<b>2.20</b>	<b>3.00</b>	<b>2.80</b>	2.80	<b>3.40</b>	<b>3.20</b>	2.60	3.00	<b>3.40</b>	2.80	<b>3.00</b>	2.60	<b>3.20</b>	65%
Case 37	<b>2.20</b>	<b>2.00</b>	<b>2.00</b>	<b>1.80</b>	<b>2.40</b>	<b>2.40</b>	<b>2.20</b>	<b>4.20</b>	<b>4.20</b>	<b>4.00</b>	<b>4.60</b>	<b>3.60</b>	<b>3.20</b>	<b>4.00</b>	<b>4.20</b>	<b>3.80</b>	<b>3.80</b>	100%
Case 38	3.00	3.00	3.20	3.40	<b>2.80</b>	3.20	3.20	<b>3.00</b>	<b>3.20</b>	2.80	<b>3.60</b>	2.60	3.00	2.80	<b>3.00</b>	2.80	2.80	29%
Case 39	<b>2.40</b>	<b>2.60</b>	<b>2.80</b>	<b>2.20</b>	3.60	3.20	<b>2.60</b>	<b>3.60</b>	<b>3.80</b>	<b>3.60</b>	<b>4.00</b>	<b>3.40</b>	2.60	<b>3.40</b>	<b>3.60</b>	<b>3.40</b>	<b>3.20</b>	82%
Case 40	<b>1.60</b>	<b>1.40</b>	<b>1.60</b>	<b>1.80</b>	<b>1.80</b>	<b>2.20</b>	<b>2.00</b>	<b>4.00</b>	<b>4.40</b>	<b>4.40</b>	<b>4.40</b>	<b>3.40</b>	<b>4.00</b>	<b>4.60</b>	<b>4.60</b>	<b>4.60</b>	<b>4.60</b>	100%
Case 41	<b>2.60</b>	<b>2.80</b>	<b>2.80</b>	<b>3.00</b>	<b>2.60</b>	<b>2.80</b>	3.00	<b>3.20</b>	<b>3.40</b>	<b>3.20</b>	<b>3.60</b>	<b>3.20</b>	<b>3.40</b>	<b>3.20</b>	<b>3.60</b>	<b>3.20</b>	<b>3.20</b>	94%
Cue 'hit' rate; honest appeals	75%	75%	69%	63%	81%	75%	81%	81%	75%	81%	69%	75%	75%	63%	75%	63%	75%	
Cue 'hit' rate; all appeals	66%	69%	69%	63%	75%	72%	75%	69%	69%	72%	69%	66%	72%	66%	75%	63%	69%	

Scores in **bold** are hits, ie case score is in expected direction (honest appeals: cue score higher than cue mean on cues to honesty, and lower than cue mean on cues to deception. Deceptive appeals: cue score higher than cue mean on cues to deception, and lower than cue mean on cues to honesty)

#### **8.3.4. Cues as Measures of Four Factors**

General inspection of the cues that discriminated between honest and deceptive appeals suggested that they could be reasonably construed as measuring four different factors: fake emotion (fake emotion, fake facial expression, and performance); norm violation (unnatural, creepy, and inversely, plausible, and normal); personal reaction (dislike, no sympathy for the appealer, and inversely, feel the appealer's pain, and feel sorry for the appealer); and genuine emotion (sad eyes, genuinely sad, genuine and heartfelt, urgency, personal and expressive voice, and vocal quiver). Given the large number of cues, therefore, it was decided to investigate the utility of collapsing the cues into these four categories. To this end, a series of partial correlations was performed. Some cues related to honesty (plausible, normal, feel the appealer's pain, and feel sorry for the appealer) were inversely scored so that in the two factors which included cues to both honesty and deception (norm violation and personal reaction), all cues related positively to deception.

##### **8.3.4.1. Factor: *fake emotion***

Biserial correlations indicated that each cue associated with the fake emotion factor (fake emotion, fake facial expression, and performance) correlated significantly with veracity in the expected direction, and all three cues correlated very strongly with each other (see Table 8.3), Cronbach's Alpha .986. Partial correlations were conducted to discover whether there was any unique variance for any of the cues in this factor. None of the three cues remained significantly correlated to deception when the other two cues were partialled out (fake emotion,  $r = .10$ ,  $p = .584$ ; fake

facial expression,  $r = .08$ ,  $p = .691$ ; performance,  $r = .03$ ,  $p = .870$ ). The scores of the three cues were, therefore, added together to create a single fake emotion cue, which significantly correlated with deception, and as well as any of the individual cues in the factor,  $r = .56$ ,  $p = .001$ .

Table 8.3: *Correlations (with significance values) for cues in 'fake emotion' factor (N = 32)*

	Fake emotion	Fake facial expression	Performance
Veracity	$r = .56$ $p = .001$	$r = .55$ $p = .001$	$r = .53$ $p = .002$
Fake emotion		$r = .97$ $p < .001$	$r = .96$ $p < .001$
Fake facial expression			$r = .95$ $p < .001$

#### 8.3.4.2. Factor: norm violation

Biserial correlations indicated that each cue associated with the norm violation factor (unnatural, creepy, and inversely, plausible, and normal) correlated significantly with veracity in the expected direction, and all four cues correlated strongly with each other (see Table 8.4), Cronbach's Alpha .906. Again, partial correlations were conducted to discover whether there was any unique variance for any of the cues in this factor. None of the four cues remained significantly correlated to deception when the other three cues were partialled out (unnatural,  $r = .06$ ,  $p = .750$ ; creepy,  $r = .15$ ,  $p = .426$ ; plausible,  $r = .11$ ,  $p = .567$ ; normal,  $r = .07$ ,  $p = .736$ ). The scores of the four cues were, therefore, added together to create a single norm violation cue, which significantly correlated with deception, and as well as any of the individual cues within the factor,  $r = .55$ ,  $p = .001$ .



Table 8.4: *Correlations (with significance values) for cues in ‘norm violation’ factor (N = 32)*

	Unnatural	Creepy	Plausible	Normal
Veracity	$r = .44$ $p = .011$	$r = .51$ $p = .003$	$r = .49$ $p = .005$	$r = .50$ $p = .004$
Unnatural		$r = .63$ $p < .001$	$r = .78$ $p < .001$	$r = .62$ $p < .001$
Creepy			$r = .68$ $p < .001$	$r = .84$ $p < .001$
Plausible				$r = .74$ $p < .001$

#### 8.3.4.3. Factor: personal reaction

Biserial correlations indicated that each cue within the personal reaction factor (dislike, no sympathy for the appealer, and inversely, feel the appealer’s pain, and feel sorry for the appealer) correlated significantly with veracity in the expected direction, and all four cues correlated very strongly with each other (see Table 8.5), Cronbach’s Alpha .978. However, none of the four cues remained significantly correlated to deception when the other three cues were partialled out (dislike,  $r = .15$ ,  $p = .428$ ; no sympathy,  $r = .18$ ,  $p = .344$ ; feel appealer’s pain,  $r = .30$ ,  $p = .110$ ; feel sorry for appealer,  $r = .19$ ,  $p = .312$ ). The scores of the four cues were, therefore, again added together to create a single personal reaction cue, which significantly correlated with deception, and equivalently to any of the individual cues in the factor,  $r = .52$ ,  $p = .002$ .

Table 8.5: *Correlations (with significance values) for cues in ‘personal reaction’ factor (N = 32)*

	Dislike	No sympathy	Feel appealer's pain	Feel sorry for appealer
Veracity	$r = .53$ $p = .002$	$r = .53$ $p = .002$	$r = .44$ $p = .011$	$r = .53$ $p = .002$
Dislike		$r = .92$ $p < .001$	$r = .90$ $p < .001$	$r = .91$ $p < .001$
No sympathy			$r = .94$ $p < .001$	$r = .94$ $p < .001$
Feel appealer's pain				$r = .94$ $p < .001$

#### 8.3.4.4. Factor: genuine emotion

Biserial correlations indicated that each cue within the genuine emotion (genuinely sad, sad eyes, genuine and heartfelt, urgency, personal and expressive voice, and vocal quiver) correlated significantly with veracity in the expected direction, and all six cues correlated strongly with each other (see Table 8.6), Cronbach's Alpha .972. However, none of the six cues remained significantly correlated with deception when the other five cues were partialled out (sad,  $r = .21$ ,  $p = .305$ ; sad eyes,  $r = .14$ ,  $p = .490$ ; genuine/heartfelt,  $r = .08$ ,  $p = .699$ ; urgency,  $r = .10$ ,  $p = .635$ ; personal and expressive voice,  $r = .07$ ,  $p = .722$ ; vocal quiver,  $r = .14$ ,  $p = .476$ ). As before then, the scores of the six cues were added together to create a single genuine emotion cue, which significantly correlated with deception, and equivalently to any of the individual cues in the factor,  $r = .50$ ,  $p = .004$ .

Table 8.6: *Correlations (with significance values) for cues in 'genuine emotion' factor (N = 32)*

Sad	Sad eyes	Genuine/h	Urgency	Personal	Vocal
-----	----------	-----------	---------	----------	-------

	earfelt				expressive voice	quiver
Veracity	$r = .52$ $p = .002$	$r = .43$ $p = .013$	$r = .50$ $p = .004$	$r = .39$ $p = .029$	$r = .46$ $p = .009$	$r = .52$ $p = .002$
Sad		$r = .91$ $p < .001$	$r = .98$ $p < .001$	$r = .79$ $p < .001$	$r = .91$ $p < .001$	$r = .94$ $p < .001$
Sad eyes			$r = .90$ $p < .001$	$r = .72$ $p < .001$	$r = .82$ $p < .001$	$r = .86$ $p < .001$
Genuine/ heartfelt				$r = .75$ $p < .001$	$r = .91$ $p < .001$	$r = .92$ $p < .001$
Urgency					$r = .76$ $p < .001$	$r = .80$ $p < .001$
Personal express. voice						$r = .89$ $p < .001$

#### 8.3.4.5. Case by case analyses using four factors

As before, case by case analyses were conducted, but this time to assess the potential of each factor to aid lie detection. Thus for each case, the score for each factor was compared with the overall factor mean. For deceptive appeals, factor scores that fell above the factor mean on the factors fake emotion, norm violation, and personal reaction, were classified as a ‘hit’, and factor scores that fell below the cue mean on the factor genuine emotion, were also classified as a ‘hit’. Similarly, for honest appeals, factor scores that fell above the factor mean on the factor genuine emotion, and below the factor mean on the other factors, were classified as a ‘hit’. In this way, a percentage of ‘hits’ was calculated for each factor (i.e. cases combined). The factors had ‘hit’ rates between 72% (norm violation) and 66% (fake emotion, personal reaction, and genuine emotion); see Table 8.7.

The percentage of ‘hits’ was also calculated for each case (i.e. factors combined). Honest appeals had ‘hit’ rates between 25 - 100%, and deceptive appeals

had ‘hit’ rates between 0 - 100% (see Table 8.7). As before, cases which scored above 50% ‘hits’ on the cues were categorised as correctly classified (the appeal scored in the expected direction on the majority of the cues), and cases that scored below 50% ‘hits’ on the cues were categorised as misclassified (the appeal did not score in the expected direction on the majority of the cues). Cases that scored 50% were categorised as unclassified. In this way, 10 of the 16 deceptive appeals were correctly classified, 10 of the honest appeals were correctly classified, and one honest appeal was unclassified. Overall, 63% of appeals were correctly classified.

Table 8.7: *Case by case analyses: factor mean, score per case (appeal), and overall ‘hit’ rate. For each case: overall ‘hit’ rate of factors. All ‘hit’s related to overall cue means*

	Fake emotion ( <i>M</i> = 8.82)	Norm violation ( <i>M</i> = 11.98)	Personal reaction ( <i>M</i> = 11.88)	Genuine emotion ( <i>M</i> = 18.36)	<b>Total hits</b>
<b><u>Deceptive Appeals</u></b>					
Case 10	<b>11.40</b>	<b>13.00</b>	<b>14.60</b>	<b>11.60</b>	100%
Case 11	<b>12.40</b>	<b>17.00</b>	<b>16.40</b>	<b>10.20</b>	100%
Case 12	<b>13.40</b>	<b>16.60</b>	<b>17.20</b>	<b>9.60</b>	100%
Case 13	8.60	11.40	11.00	18.60	0%
Case 14	8.20	<b>11.60</b>	<b>12.60</b>	<b>15.20</b>	75%
Case 15	<b>9.40</b>	<b>14.40</b>	<b>12.60</b>	<b>16.00</b>	100%
Case 16	8.00	<b>14.20</b>	10.60	22.60	25%
Case 17	<b>12.66</b>	<b>15.33</b>	<b>16.99</b>	<b>10.01</b>	100%
Case 18	<b>11.20</b>	<b>13.40</b>	<b>14.00</b>	<b>15.40</b>	100%

Case 19	<b>9.00</b>	10.80	11.80	18.80	25%
Case 20	7.80	10.40	10.20	21.40	0%
	Fake emotion ( <i>M</i> = 8.82)	Norm violation ( <i>M</i> = 11.98)	Personal reaction ( <i>M</i> = 11.88)	Genuine emotion ( <i>M</i> = 18.36)	<b>Total hits</b>
Case 21	7.60	10.40	10.40	20.40	0%
Case 22	6.20	9.20	8.20	25.20	0%
Case 23	<b>12.40</b>	<b>14.80</b>	<b>15.60</b>	<b>13.20</b>	100%
Case 24	<b>11.20</b>	<b>14.00</b>	<b>14.20</b>	<b>14.00</b>	100%
Case 25	<b>11.40</b>	<b>15.20</b>	<b>16.00</b>	<b>14.20</b>	100%
<b>Factor ‘hit’ rate; deceptive appeals</b>	63%	69%	63%	63%	
<b><u>Honest Appeals</u></b>					
Case 26	9.40	<b>11.20</b>	<b>11.00</b>	16.60	50%
Case 27	<b>8.40</b>	<b>10.60</b>	<b>11.60</b>	17.00	75%
Case 28	<b>8.20</b>	<b>10.20</b>	<b>10.80</b>	<b>19.40</b>	100%
Case 29	<b>5.40</b>	<b>9.40</b>	<b>8.40</b>	<b>26.20</b>	100%
Case 30	<b>8.00</b>	12.00	<b>10.20</b>	<b>20.60</b>	75%
Case 31	9.20	<b>11.20</b>	12.20	16.80	25%
Case 32	<b>5.60</b>	<b>8.60</b>	<b>7.60</b>	<b>25.60</b>	100%
Case 33	<b>6.00</b>	<b>8.80</b>	<b>8.40</b>	<b>22.00</b>	100%
Case 34	9.00	13.60	12.00	<b>20.20</b>	25%
Case 35	<b>7.20</b>	12.20	12.20	16.80	25%
Case 36	9.00	<b>10.60</b>	12.00	17.80	25%
Case 37	<b>6.20</b>	<b>9.40</b>	<b>9.00</b>	<b>25.20</b>	100%
Case 38	9.20	12.60	12.80	<b>18.40</b>	25%
Case 39	<b>7.80</b>	<b>11.80</b>	<b>11.20</b>	<b>22.00</b>	100%
Case 40	<b>4.60</b>	<b>8.20</b>	<b>7.00</b>	<b>26.40</b>	100%
Case 41	<b>8.20</b>	<b>11.00</b>	<b>11.40</b>	<b>20.20</b>	100%
<b>Factor ‘hit’ rate; honest appeals</b>	69%	75%	69%	69%	
<b>Factor ‘hit’</b>					

<b>rate; all appeals</b>	66%	72%	66%	66%
------------------------------	-----	-----	-----	-----

Scores in **bold** are hits, i.e. case score is in expected direction.

### 8.3.5. Accuracy of Observers' Credibility Judgements

To investigate the hypothesis that observers would produce accuracy rates in detecting deception above chance, overall accuracy rates were calculated for each observer. Any cases which were known to an observer were not included in the accuracy scores. Mean observer accuracy was 67%. For each observer, the percentage of appeals correctly classified on veracity was compared with the percentage of appeals incorrectly judged on veracity, using a Wilcoxon test. Observers correctly classified appeals (median percentage = 63) significantly more often than they incorrectly classified appeals (median percentage = 37) ( $Z = 2.02$ ,  $p = .043$ ).

Analyses were then conducted to investigate the hypothesis that observers would not demonstrate a truth bias. For each observer, the percentage of deceptive appeals correctly classified was compared with the percentage of deceptive appeals not correctly classified using a Wilcoxon test. Observers correctly classified deceptive appeals (median percentage = 60) significantly more often than they incorrectly classified deceptive appeals (median percentage = 40), ( $Z = 2.03$ ,  $p = .042$ ), suggesting that observers were able to accurately classify deceptive appeals above the level of chance. Again, for each observer, the percentage of honest appeals correctly classified was compared with the percentage of honest appeals not correctly classified using a Wilcoxon test. Observers correctly classified honest appeals

(median percentage = 63) significantly more often than they incorrectly classified honest appeals (median percentage = 37), ( $Z = 2.03$ ,  $p=.042$ ), suggesting that observers were able to accurately classify honest appeals above the level of chance. Finally, the percentage of deceptive appeals correctly classified was compared with the percentage of honest appeals correctly classified using a Wilcoxon test. The difference between observers' correct classifications of deceptive appeals (median percentage = 60), and honest appeals (median percentage = 63), was not significant ( $Z = 0.73$ ,  $p = .465$ ), suggesting that there was no difference in the accuracy with which observers were able to classify liars and truth-tellers.

#### **8.4 Discussion**

The findings of Study 4 have a number of possibly noteworthy implications. For example, they provide further support for the general idea that investigating the cues used by people known to be accurate at detecting deception may be a useful methodological approach. As previously noted, this methodology is nascent in deception research (Mann et al., 2006; O'Sullivan & Ekman, 2004); nevertheless, a number of cues identified by accurate participants in Study 1 successfully discriminated between a different sample of deceptive and honest appeals, and all with substantial effect sizes. This lends actuarial validity to the cues, suggesting that although they were not generated from existing theoretical approaches, but rather using a data-driven, bottom-up approach, they may have utility in the context of appeals for help with missing or murdered relatives.

Furthermore, the majority of the cues found to discriminate between honest and deceptive appeals were previously unidentified cues. This endorses the view that ecologically valid, high stakes deception situations may elicit context-specific cues to deception not apparent in typical low stakes, laboratory situations. For example, as originally predicted in Study 2, emotion was a key element in many of the cues which discriminated between deceptive and honest appeals. Hence, deceptive appealers were rated higher than honest appealers on cues related to fake emotion (fake emotion, fake facial expression, and putting on a performance). This suggests not only that simulated emotion is a useful cue to deception in the context of making an appeal, but also that observers are able to detect simulated emotion in this context. Similarly, honest appealers were rated higher than deceptive appealers on a variety of general and specific cues related to emotional authenticity (such as, genuinely sad, genuine/heartfelt appeal, urgency, personal and expressive voice, voice quivering with emotion, and sad eyes). The vocal cues (personal and expressive voice, and voice quivering with emotion) are particularly interesting, as for decades efforts have been made to develop voice analysis software to detect deception, using measures such as fundamental frequency, jitter, shimmer, and intensity. However, previous attempts to use measures of vocal characteristics as indicators of deception have generally been without success, and have consistently demonstrated that the accuracy of such cues in detecting deception are around the levels of chance (see, for example, Eriksson & Lacerda, 2007; Harnsberger, Hollien, Martin & Hollien, 2009; Giddens et al., 2013; Horvath, McCloughlan, Weatherman & Slowik, 2013). In contrast, the present results suggest that subjective impressions of vocal characteristics may capture an aspect of deceptive behaviour which it has not been possible to measure more objectively.



It seems likely that a number of the cues identified in the present study may be specific to the context of making an appeal; hence, to what extent these cues are context specific, or may have utility in different high stakes situations, is a question for further research. However, the underlying factors to which the cues may be related, such as, emotional authenticity (including fake emotion, and genuine emotion), norm violation, and the personal reaction elicited in observers, may have relevance across particular high stakes situations, and may also point to areas in which theory could be developed. Importantly also, many of the previously unidentified but discriminatory cues, are not readily accommodated by existing theoretical approaches; this suggests that the development of new approaches may be beneficial when considering real world, high stakes deceptive behaviour. This issue is addressed more fully in the general discussion.

The accuracy of classification of appeals as honest or deceptive achieved in the case by case analyses (69%), was lower than in Study 2 (78%), and collapsing the cues in to four factors did not increase classification accuracy (65%). It is notable, however, that the 'hit' rate for some individual cues was high; the cues creepy, no sympathy for the appealer, and voice quivering with emotion, achieved classification accuracy of 75%, and the cues dislike the appealer, genuine and heartfelt, and normal, achieved classification accuracy of 72%. This suggests that these cues in particular may have utility in discriminating between honest and deceptive appeals.

As hypothesised, observer accuracy in detecting deception was higher than chance (67%), and truth bias was eliminated, supporting the similar findings from Study 2. As previously discussed, this suggests that in real life, high stakes lies, cues to deception may be more prominent and therefore easier for observers to discern, than in typical, low stakes, laboratory experiments. Again, as in Study 2, the sample

size of observers in the present study was too small to draw firm conclusions (although the sample size of ‘senders’ was much larger), and observer accuracy will be more fully explored in Study 7.

To summarise, the present study suggests that, in the high stakes context of appeals at least, untrained observers can show some success in differentiating between truth and lies using simple subjective judgements if given a range of specific criteria. Moreover, a number of the cues to deception that emerged as useful in the present studies had not previously been identified; it appears that use of high stakes lies, and a between subjects design, may produce new cues to deception, and this clearly has implications for our understanding of deceptive behaviour. However, the extent to which these are specific to the context of making an appeal for help with a missing or murdered relative remains unclear. Another issue that requires further investigation is the extent to which some of the cues may reflect, as yet unidentified, aberrant characteristics of the kinds of individuals who commit the crimes that the deceptive appellars have committed, and then lie about it, rather than their acts of deception per se. In the present situations, and with most other high stakes deception contexts, honest and deceptive people differed not only in terms of whether they were lying, but also whether they had committed serious crimes. Any related between groups individual differences will obviously be absent from standard low stakes laboratory situations where participants are randomly assigned to deceptive and truthful conditions. The methodological and theoretical implications of this are addressed in the general discussion.

Having considered subjective ratings, in the next chapter we move once again to more objective frequency counts.

# **Chapter 9**

## **Study 5: Frequency Counts of Verbal Cues to Deception in Appeals**

### **9.1 Introduction**

The main aim of Study 5, described in this chapter, was to replicate and extend the verbal frequency count findings of Study 3 on a new and larger sample set. As noted previously, for applied purposes, verbal cues to deception may have particular utility for two reasons: first, they are amenable to frequency counts and, therefore, may be more objectively measured than, for example, more subjective observer impressions, such as those investigated in Study 4. Second, verbal cues are available in several recording media, so that if video of the communication under scrutiny is not available, audio recordings and also transcripts can provide the necessary data.

As noted in Chapter 2, when a full account from the ‘sender’ is available, established verbal analysis techniques such as Reality Monitoring and Content Based Criteria Analysis can be applied; these have been found to produce reasonably high accuracy in discriminating between truth and deception, at around 72% - 76% (Masip et al., 2005; Vrij, 2005). Nevertheless, there are circumstances in which the material required for these techniques is not available because there is little or no account to

analyse. This is often the case with appeals for help with missing or murdered relatives' i.e. an appeal does not necessarily contain an account of an event. Nevertheless, some of the findings of this thesis so far suggest that, at least in some contexts, one may still be able to identify useful verbal cues to deception when very limited information is available. Thus all but one of the verbal cues (focused on the relative) identified in Study 1 were found to differ between honest and deceptive appeals either significantly, or with large effect sizes, when assessed with frequency counts using the same set of stimulus materials in Study 3. Hence, as mentioned, the main aim of Study 5 was to attempt to replicate and extend these findings. However, a second aim of Study 5 was to investigate which verbal cues to deception identified in research in other high stakes contexts, would also have utility in the context of appeals. Specifically, it was hypothesised that verbal cues which have differentiated between honesty and deception in other high stakes contexts, (equivocation, repetition, speech errors, filled pauses), would also discriminate between deceptive and honest appeals.

## **9.2. Method**

### **9.2.1. *Materials and Procedure***

The 32 audio recordings of the honest and deceptive appeals for help with missing or murdered relatives used in Study 4 were transcribed; transcripts varied in length between 15 and 336 words ( $M = 110.81$ ,  $SD = 80.64$ ). See Appendix 8 for transcripts of the appeals.

Following protocols established in previous research on high stakes deception (Mann et al., 2002; ten Brinke & Porter, 2012), one coder counted frequencies of predetermined verbal behaviours in the transcripts of the video clips, and a second coder, who was blind to the veracity of the appeals, coded a random sample of 25% of the deceptive appeals and 25% of the honest appeals. Both coders were graduate students studying deception. Inter-rater reliabilities for the frequency counts were high, ranging from  $r = .90$  to  $r = 1.00$ , with 13 of the 18 cues having  $r = 1.00$  (see Table 9.1). In two cases, inter-rater reliability could not be computed because at least one variable was constant (due to floor effects). As the inter-rater reliability for the frequency counts was very high for the random sample, frequency counts produced by the coder who had coded all the transcripts were used in all analyses.

The verbal behaviours included in the study were: cues originally identified in Study 1, and which in Study 3 had been found to discriminate between honest and deceptive appeals, or to have had large effect sizes ( $d > 0.75$ ) in the expected direction (again, therefore, actuarial rather than theoretical prediction provided the primary rationale for the inclusion of these cues at this stage); cues identified in other research on high stakes deception; and two cues noticed by the researcher.

Verbal behaviours that were expected to be related to honesty in the context of an appeal were: positive emotion expressed towards the relative (including expressions of love, positive descriptions, and terms of endearment); avoiding the use of brutal language (for example, saying ‘gone’ instead of ‘murdered’); and expressions of hope (relevant only in cases where the body of the relative had not been found), all of which were derived from Study 1. A plea for help has been found to relate to honesty in Harpster et al.’s (2009), study of 911 homicide calls, and it was included in the present study, as it was considered that this behaviour may have

particular relevance to the context of appeals. A further two possible cues were noticed by the researcher whilst transcribing the appeals and were included in the analysis; they were, expressions of concern for the relative or expressions of grief or pain (for example, 'we just want to know she's all right' and 'this has distressed all of our family'); and references to norms of emotion or behaviour, including violations of norms (for example 'how could anybody do this' and 'as any parent knows'). These two cues were expected to be related to honesty.

Cues derived from Study 1, and that were expected to be related to deception in the context of an appeal were: sentence or statement that does not make sense (illogical or unclear); avoiding first person pronouns (for example 'Just want her back'); referring to other people or groups of people; expressions indicating a lack of hope (relevant only in cases where the body of the relative had not been found); lists (for example, of people, actions, emotions); irrelevant statements; equivocation (including words or phrases that minimise, qualify and convey uncertainty or vagueness, for example 'just', 'kind of', 'possibly', 'a bit', 'don't know'); and the use of brutal language or detail about the relative (for example 'blood' and 'murdered'). A number of these cues have also been found to relate to deception in other high stakes research: thus Harpster et al.'s (2009) analysis of 911 homicide calls found that acceptance of death, extraneous information, and conflicting facts predicted deception; acceptance of death could be considered as equivalent to lack of hope, extraneous information as equivalent to irrelevant statements, and conflicting facts as equivalent to sentence or statement that does not make sense. Equivocation has also been found to relate to deception across other high stakes contexts, as in both written statements made to police (Adams & Jarvis, 2006), and also in appeals (ten Brinke & Porter, 2012). Cues included that were derived entirely from previous

research on high stakes deception were: phrase repetition (Davis et al., 2005; Harpster & Adams, 2008), filled pauses (Davis et al., 2005), and speech errors (Davis et al., 2005; Vrij & Mann, 2001a). Speech errors were included because, although this cue did not reach significance or have a large effect size in Study 3, it has strong theoretical a priori support as an indicator of cognitive load (Sporer & Schwandt, 2006).

For a full description of each verbal cue, and instructions to coders for frequency counts, see Appendix 9.

### **9.3. Results**

#### **9.3.1. *Analyses of Cues***

Preliminary analyses revealed no effect of the word count of the appeals on veracity,  $F(1,31) = 0.01$ ,  $p = .990$ , so this was not included as a variable in the following analyses. However, to account for the differing lengths of the appeals, each cue score for each appeal was divided by the word count of the appeal, to produce a frequency count for each cue as a proportion of the total appeal. As this produced very small numbers, for the purposes of clarity this proportion was then multiplied by 100, which produced a frequency count per 100 words.

To investigate which cues were associated with honesty, a MANOVA with veracity as a between-subjects independent variable, was conducted on the adjusted frequency counts for the following cues: positive emotion expressed towards the relative, avoid brutal language/detail, makes a plea, expressions of concern/pain,

references to norms. Means, standard deviations, confidence intervals,  $F$  ratios,  $p$  values and effect sizes are presented in Table 9.1.

Using Pillai's Trace, the multivariate test was significant,  $F(1,26) = 2.88, p = .033$ . Follow-up univariate analyses yielded significant effects for the cue reference to norms, with means in the expected direction. The cue expressions of concern, grief, or pain, also approached significance ( $p = .032$ , one-tailed); again, as expected, honest appeals contained a larger proportion of expressions of concern, grief, or pain than deceptive appeals.

The cue avoid brutal language or detail also approached significance ( $p = .038$ , one-tailed). However, it was assumed that this cue might be affected by the status of the appellant's relative; i.e. appellants would be more likely to use this cue more frequently in cases where the relative was publicly known to be dead. Accordingly, a  $2 \times 2$  ANOVA (veracity  $\times$  status) was conducted on the data. This showed significant effects for veracity ( $F(1,28) = 5.52, p = .042$ ) and status ( $F(1,28) = 8.21, p = .008$ ). As expected, appeals for help with dead relatives ( $M = 1.52, SD = 1.66$ ) contained a larger proportion of avoidance of brutal language than appeals for help with missing relatives ( $M = 0.32, SD = 0.81$ ). The interaction between veracity and status was not significant ( $F < 1$ ).

Levene's test for homogeneity of variance was significant for the cues references to norms, expressions of concern, grief, or pain, and avoid brutal language or detail. Welch's Tests were run on these cues as a check, and produced results comparable to the univariate tests (for references to norms,  $p = .039$ , expressions of concern, grief or pain,  $p = .067$ , and avoids brutal language or detail,  $p = .076$ ).

As the cue expressions of hope would not be relevant in cases where the relative was publicly known to be dead, a separate analysis was conducted for this



cue using only cases in which the relative was missing (i.e. no body had been found). Ten deceptive appeals and 10 honest appeals were included in this analysis. A one way ANOVA (with veracity as the factor) showed, as expected, that honest appeals ( $M = 2.27$ ,  $SD = 1.91$ ) contained a significantly higher proportion of expressions of hope than deceptive appeals ( $M = 0.63$ ,  $SD = 0.94$ ),  $F(1,19) = 5.96$ ,  $p = .025$ .

Although the cue 'positive emotion expressed towards the relative' did not discriminate significantly using the mean percentage scores (see Table 9.1), examination of the data showed that an extreme outlier was producing a very pronounced skew in the data (one deceptive appealer produced a very high number of expressions of positive emotion towards his fiancé). Consequently, this cue was also coded dichotomously as present/absent, and a further binary analysis was conducted. This yielded a significant result in the predicted direction; expressions of positive emotion towards the relative were more likely to be present in honest appeals ( $n = 11$  or 69%), than in deceptive appeals ( $n = 5$  or 31%),  $\chi^2(1) = 4.50$ ,  $p = .034$ .

To investigate which cues were associated with deception, two separate MANOVAs, with veracity as a between subjects independent variable, were conducted on the percentage frequency scores of the cues. One was conducted on the cues derived originally from Study 1 which focused specifically on appeals (sentence or statement that does not make sense, avoids using first person pronoun, referring to others, lists, irrelevant statements, and using brutal language or detail), and the other was conducted on four cues that have been found to relate to deception across contexts in previous research on high stakes lies (equivocation, speech errors, phrase repetition, and filled pauses).

Using Pillai's Trace, the multivariate test on the cues identified in previous research on high stakes lies was not significant,  $F(1,27) = 1.63, p = .196$ . However, it can be noted that all the cues had means in the expected direction, and two of the cues, equivocation and speech errors, had substantial effect sizes, particularly the former (see Table 9.1). Consequently, notwithstanding the lack of significance of the overall MANOVA, the cues equivocation and speech errors were also analyzed individually; deceptive appeals contained significantly more equivocal language than honest appeals,  $F(1,31) = 4.79, p = .036$ , and also more speech errors, although this difference was significant only on a one-tailed test,  $F(1,31) = 3.70, p = .064$ .

The multivariate test on the cues originally derived from Study 1, sentence or statement that does not make sense, avoids using first person pronoun, referring to others, lists, irrelevant statements, and using brutal language or detail, was not significant,  $F(1,25) = 1.06, p = .411$ .

As the cue 'lack of hope of finding the relative alive', would not be relevant in cases where the relative was publicly known to be dead, this was not included in the MANOVA; instead a separate analysis was conducted for this cue using only cases in which the relative was missing (i.e. no body had been found). However, a one way ANOVA (with veracity as the factor) was not significant,  $F(1,19) = 1.00, p = .331$ .

Table 9.1: *Inter-rater reliability, percentage means (SDs), CIs, F ratios and p values, and effect sizes for verbal cues*

Cue	Inter-rater reliability ( <i>r</i> )	<u>Deceptive</u> Mean ( <i>SD</i> ) CI	<u>Honest</u> Mean ( <i>SD</i> ) CI	<i>F</i> ratio ( <i>p</i> )	Effect size
<u>Verbal cues to honesty</u>					

Positive emotion towards relative <sup>a</sup>	$r = 1$ $p < .001$	1.90 (3.77) 0.13 – 3.68	3.70 (3.15) 1.93 – 5.47	$F(1,31) = 2.14$ $p = .154$	$d = 0.52$
Avoids brutal language/detail*	$r = 1$ $p < .001$	0.36 (0.96) -0.29 – 1.00	1.19 (1.51) 0.54 – 1.83	$F(1,31) = 3.43$ $p = .074$	$d = 0.65$
Expressions of hope**	$r = 1$ $p < .001$	0.63 (0.94) -0.047 – 1.301	1.91 (0.60) 0.904 – 3.630	$F(1,19) = 5.96$ $p = .025$	$d = 1.62$
Makes a plea	$r = 1$ $p < .001$	3.77 (4.52) 1.87 – 5.67	2.69 (2.68) 0.79 – 4.58	$F(1,31) = 0.68$ $p = .417$	$d = 0.29$
Expressions of concern /pain*	$r = .90$ $p = .003$	1.47 (2.11) -0.24 – 3.17	3.74 (4.22) 2.04 – 5.45	$F(1,31) = 3.71$ $p = .063$	$d = 0.68$
References to norms of emotion or behaviour**	$r = 1$ $p < .001$	0.08 (0.24) -0.50 – 0.65	0.98 (1.58) 0.40 – 1.55	$F(1,31) = 5.09$ $p = .032$	$d = 0.78$

**Verbal cues to deception**

Equivocation**	$r = 1$ $p < .001$	4.69 (4.40) 2.93 – 6.45	2.03 (2.07) 0.27 – 3.78	$F(1,31) = 4.79$ $p = .036$	$d = 0.77$
Speech errors*	$r = .98$ $p < .001$	4.90 (3.75) 3.39 – 6.41	2.89 (1.85) 1.38 – 4.40	$F(1,31) = 3.70$ $p = .064$	$d = 0.68$
Phrase repetition	$r = .98$ $p < .001$	4.84 (4.92) 2.71 – 6.97	4.72 (3.25) 2.59 – 6.84	$F(1,31) = 0.01$ $p = .934$	$d = 0.03$

Cue	Inter-rater reliability ( $r$ )	Deceptive Mean ( $SD$ ) CI	Honest Mean ( $SD$ ) CI	$F$ ratio ( $p$ )	Effect size
Filled pauses	$r = 1$ $p < .001$	1.47 (1.79) 0.60 – 2.34	1.10 (1.62) 0.23 – 1.97	$F(1,31) = 0.38$ $p = .542$	$d = 0.22$
Does not make sense	<i>All cases checked</i> <i>0</i>	0.68 (1.07) 0.24 – 1.13	0.42 (0.61) -0.03 – 0.86	$F(1,31) = 0.75$ $p = .393$	$d = 0.31$
Avoids using 1 <sup>st</sup> person pronoun	$r = 1$ $p < .001$	0.81 (1.32) 0.07 – 1.56	0.82 (1.59) 0.08 – 1.57	$F(1,31) = 0.00$ $p = .981$	$d = 0.01$
Refers to others	$r = 1$ $p < .001$	1.39 (1.75) 0.39 – 2.38	2.13 (2.13) 1.13 – 3.12	$F(1,31) = 1.16$ $p = .291$	$d = 0.38$
Lack of hope	<i>All cases checked</i> <i>0</i>	0.67 (2.11) -.32 – 1.66	0.00 (0.00) -0.99 – 0.99	$F(1,19) = 1.00$ $p = .331$	$d = 0.45$
Lists	$r = 1$ $p < .001$	0.54 (0.89) 0.11 – 0.96	0.10 (0.77) 0.57 – 1.42	$F(1,31) = 2.41$ $p = .131$	$d = 0.55$
Irrelevant statements	$r = 1$ $p < .001$	0.49 (0.90) 0.07 – 0.91	0.22 (0.76) -0.21 – 0.64	$F(1,31) = 0.86$ $p = .362$	$d = 0.33$

Uses brutal language/detail	$r = 1$ $p < .001$	0.38 (0.95) 0.03 – 0.73	0.06 (0.19) -0.29 – 0.42	$F(1,31) = 1.68$ $p = .204$	$d = 0.46$
--------------------------------	-----------------------	----------------------------	-----------------------------	--------------------------------	------------

*Note:* <sup>a</sup> Cue differs significantly between deceptive/honest appeals when coded dichotomously as present/absent. <sup>\*\*</sup> Means differ significantly from each other on a two-tailed test. <sup>\*</sup> Means differ significantly from each other on a one-tailed test

### 9.3.2. Case by Case Analyses

As in Studies 2, 3 and 4, case by case analyses were conducted to assess the potential of each discriminatory cue to predict veracity, and also to combine information from the discriminatory cues to predict the veracity of the individual cases (appeals). The cues considered in this analysis were those that had significantly discriminated between liars and truth-tellers on a two-tailed test (equivocation, references to norms of emotions/behaviour, expressions of hope, and expressions of love/positive descriptions of the relative as a dichotomous variable). For each case, the adjusted frequency score for each cue was compared with the overall cue mean. For the deceptive appeals, cue scores that fell above the cue mean on the cue to deception (equivocation) were classified as a ‘hit’, and cue scores that fell below the cue mean on cues to honesty (expressions of hope, and references to norms of emotion or behaviour) were also classified as a ‘hit’. For the honest appeals, cue scores that fell above the cue mean on cues to honesty were classified as a ‘hit’, along with cue scores that fell below the cue mean on the cue to deception. The cue ‘expressions of positive emotion towards the relative’ (which related to honesty when coded dichotomously as present/absent), was classified as a ‘hit’ if present in an honest appeal, and a ‘hit’ if absent in a deceptive appeal. In this way, a percentage of ‘hits’ was calculated for each cue (i.e. the percentage number of cases in which a ‘hit’ was

scored). The individual cues had ‘hit’ rates between 65% and 69%. See Table 9.2 for details.

As in previous studies in the present thesis, the information from the cues was then combined to classify individual cases as deceptive or honest. For each case, the number of ‘hits’ on the cues was calculated (i.e. the number of times the appeal scored in the expected direction on each of the cues). Honest appeals had hit rates between 25% and 100% ( $M = 65\%$ ), deceptive appeals had hit rates between 25% and 100% ( $M = 73\%$ ), and the mean hit rate was 69% (see Table 9.2). Cases which scored above 50% ‘hits’ on the cues were categorised as correctly classified (the appeal scored in the expected direction on the majority of the cues). Cases which scored 50% ‘hits’ on the cues were categorised as unclassified, and cases that scored below 50% ‘hits’ on the cues were categorised as misclassified (the appeal did not score in the expected direction on the majority of the cues). In this way, 13 deceptive appeals were correctly classified (81%), 2 were unclassified (13%) and 1 was misclassified (6%). Of the honest appeals, 9 were correctly classified (56%), 4 were unclassified (25%) and 3 were misclassified (19%). Overall, therefore, using these criteria, 69% of cases were correctly classified, 19% were unclassified and 13% were misclassified.

Table 9.2: *Case by case analyses using verbal cues: score per case (appeal), and overall ‘hit’ rate. For each case: overall ‘hit’ rate of cues. All ‘hit’s related to overall cue means*

	Norms ( $M = .53$ )	Positive emotion; dichotomous	Hope ( $M = 1.45$ )	Equivocation ( $M = 3.36$ )	Total hits
<b><u>Deceptive appeals</u></b>					
Case 10	<b>0.00</b>	<b>Absent</b>	1.88	<b>5.26</b>	75%

Case 11	<b>0.00</b>	<b>Absent</b>	n/a	0.00	67%
Case 12	<b>0.00</b>	<b>Absent</b>	n/a	0.97	67%
Case 13	<b>0.00</b>	Present	<b>0.00</b>	<b>12.00</b>	75%
Case 14	<b>0.00</b>	Present	n/a	<b>4.17</b>	67%
Case 15	<b>0.00</b>	<b>Absent</b>	n/a	1.03	67%
Case 16	<b>0.00</b>	<b>Absent</b>	n/a	<b>7.41</b>	100%
Case 17	<b>0.00</b>	<b>Absent</b>	<b>0.00</b>	<b>14.29</b>	100%
Case 18	<b>0.00</b>	Present	<b>0.00</b>	0.00	50%
Case 19	<b>0.00</b>	<b>Absent</b>	<b>0.00</b>	<b>6.90</b>	100%
Case 20	<b>0.00</b>	<b>Absent</b>	<b>0.00</b>	0.00	75%
Case 21	<b>0.00</b>	<b>Absent</b>	2.44	<b>4.88</b>	75%
Case 22	<b>0.00</b>	Present	<b>0.00</b>	1.87	50%
Case 23	0.91	Present	<b>0.46</b>	1.37	25%
Case 24	<b>0.00</b>	<b>Absent</b>	1.49	<b>8.96</b>	75%
Case 25	<b>0.30</b>	<b>Absent</b>	n/a	<b>5.95</b>	100%
<hr/>					
<b>Cue 'hit rate: deceptive appeals</b>	94%	69%	70%	56%	
<hr/>					
<b><u>Honest appeals</u></b>					
Case 26	0.00	Absent	<b>2.12</b>	<b>2.11</b>	50%
Case 27	0.00	Absent	<b>4.76</b>	7.94	25%
Case 28	<b>2.53</b>	<b>Present</b>	1.27	<b>2.53</b>	75%
Case 29	<b>0.70</b>	<b>Present</b>	n/a	<b>0.00</b>	100%
Case 30	<b>3.23</b>	Absent	n/a	<b>1.61</b>	67%
Case 31	0.00	<b>Present</b>	1.18	3.53	25%
<hr/>					
	Norms	Positive emotion; dichotomous	Hope	Equivocation	Total hits
	( <i>M</i> = .53)		( <i>M</i> = 1.45)	( <i>M</i> = 3.36 )	
<hr/>					
Case 32	<b>0.83</b>	<b>Present</b>	n/a	<b>1.65</b>	100%
Case 33	0.00	<b>Present</b>	n/a	<b>3.03</b>	67%
Case 34	<b>5.56</b>	Absent	n/a	4.17	33%
Case 35	0.00	Absent	<b>2.27</b>	<b>0.00</b>	50%

Case 36	<b>1.36</b>	<b>Present</b>	n/a	<b>1.36</b>	100%
Case 37	0.00	<b>Present</b>	0.00	<b>0.00</b>	50%
Case 38	0.00	<b>Present</b>	0.00	<b>2.33</b>	50%
Case 39	0.00	<b>Present</b>	<b>3.61</b>	<b>0.00</b>	75%
Case 40	<b>1.43</b>	<b>Present</b>	<b>5.71</b>	<b>0.00</b>	100%
Case 41	0.00	<b>Present</b>	<b>1.76</b>	<b>2.20</b>	75%
<hr/>					
<b>Cue ‘hit rate: honest appeals</b>	44%	69%	60%	81%	
<hr/>					
<b>TOTAL HITS:</b>	69%	69%	65%	69%	
<hr/>					

Scores in **bold** are ‘hits’, i.e. case score is in expected direction.

#### 9.4. Discussion

There are several possibly noteworthy implications of the findings from the present study. First, it appears that, as hypothesised, there are some verbal cues that have utility in discriminating between honesty and deception across high stakes situations. Equivocation was associated with deception in the present study, and has also been related to deception in an earlier study of appeals (ten Brinke & Porter, 2012), in a study of written statements made to police (Adams & Jarvis, 2006), and also in a meta-analysis (DePaulo et al., 2003). As previously noted, this cue has strong theoretical underpinnings; researchers have suggested that deception may be associated with linguistic constructions that inhibit direct communication, and distance the liar from the subject of their speech, and equivocal language could thus be construed as a form of psychological evasion or distancing from feelings of guilt (Zuckerman et al., 1981). In the specific context of appeals, the use of equivocal language may be an attempt to evade or diminish the psychological conflict produced by the discrepancy between the deceptive appealers’ knowledge of their own guilt, and their appeals for help.

The findings regarding the relationship between speech errors and deceptive communication in the present study were not robust, and should be treated with caution. Nevertheless, it is worth noting that the trend for deceptive appealers to produce more speech errors than honest appealers, is in line with other research on high stakes deception in different contexts; speech errors have been previously associated with deception in statements to assistant district attorneys (Davis et al., 2005), and in a police suspect interview (Vrij & Mann, 2001). As with equivocation, this cue also has strong theoretical underpinnings; as previously noted, it is widely accepted that lying is more cognitively demanding than telling the truth (see, for example, DePaulo et al., 2003; Frank & Feeley, 2003; Vrij et al., 2008; Zuckerman et al., 1981), and speech errors may be a manifestation of this increase in cognitive load.

The second noteworthy implication of the present findings is that five previously unidentified, and possibly context-specific, verbal cues discriminated between deceptive and honest appeals. Although clearly these findings need replication before firm conclusions can be drawn, they support the hypothesis that some context-specific cues may discriminate between honest and deceptive appeals. References to norms of emotion/behaviour were significantly more prevalent in honest appeals than in deceptive appeals, and in cases in which a body had not been found, honest appeals contained significantly more expressions of hope of finding the missing relative alive than did deceptive appeals. Honest appeals were also significantly more likely to include an expression of positive emotion towards the relative than deceptive appeals. Furthermore, honest appeals contained marginally significantly more expressions of concern, grief or pain, and avoidances of the use of brutal language or detail.



As expected, in the context of an emotionally-charged event such as making an appeal for a missing or murdered relative, context-specific, verbal cues related to emotion were important. Theoretical approaches to cues to deception have traditionally focused on emotions that may be generated by an act of deception, for example fear, anxiety, guilt, shame and excitement (Ekman, 1992; Zuckerman et al., 1981); however, the present findings endorse the view that, in high stakes contexts, such as that of making an appeal, the patterns of emotion may be more complex and context specific, requiring a different kind of explanation. For example, in appeals, the lack of an appropriate or expected emotion may also be important; the newly-identified cues all related to honesty, and none of the verbal cues derived originally from Study 1, and expected to be related to deception, were found to discriminate between honest and deceptive appeals in this larger sample. This suggests that it may be the lack of honest behaviours, rather than the production of deceptive behaviours, which is particularly salient in this context. So, for example, it could be expected that somebody making an honest appeal may be more likely to express positive emotion towards their relative, whereas somebody who has recently killed their relative may lack positive emotion towards the relative. Similarly, it could be expected that somebody making an honest appeal may produce more expressions of concern, grief or pain about the disappearance or death of the relative, than somebody who has recently killed their relative and may lack these emotions. Finally, it could be expected that whereas an honest appealer may avoid using brutal language or detail about their relative, possibly due to an unwillingness to express any brutality towards their relative, a deceptive appealer may not have the same reluctance.

In cases in which the relative was not publicly known to be dead, honest appealers were found to produce more expressions of hope of finding their relative

alive than deceptive appealers, and this may be a cognitively driven behaviour: the deceptive appealers (except in one case where the relative was kidnapped but still alive) had knowledge that their relative was already dead and would never be found alive, and this may account for why deceptive appealers rarely expressed hope of finding their relative alive. Furthermore, honest appeals contained more references to normal emotion and behaviour than deceptive appeals, and this may also be in part a cognitively driven behaviour: honest appealers asking ‘how could somebody do this?’ or ‘why would anybody want to hurt her?’ may be asking a genuine question to which they do not know the answer, whereas these would not be questions for a deceptive appealers, who knows how and why somebody could hurt their relative. Speculatively, this behaviour may also be related to cognitive dissonance or some types of personality disorder: in an attempt to mitigate cognitive dissonance, deceptive appealers may be less likely to refer to norms of emotions and behaviours which they know that they have violated; and in some cases it may be that the deceptive appealers does not have an internal representation of those norms and so does not refer to them.

It is clear that the newly-identified, and likely context-specific cues, are difficult to explain in terms of any single, overarching, existing theoretical perspective; it may be that new approaches to deceptive behaviour in high stakes contexts are required. The possible theoretical implications of this are discussed more fully in the general discussion.

Finally, the prediction that clusters of behaviours would be useful in predicting veracity was supported. The case by case analyses, which combined information from the cues to predict the veracity of individual cases, correctly classified 69% of the appeals. Furthermore, careful examination of the case by case

analyses suggests that it is important to consider the results specifically in terms of real world applications. For example, the cue ‘references to norms of emotion or behaviour’ had an overall ‘hit’ rate of 69% in discriminating between honest and deceptive appeals; however, this cue had a 94% ‘hit’ rate for deceptive appeals, and only a 44% ‘hit’ rate for honest appeals. In terms of applicability, this suggests that an appeal that contains references to norms is very likely to be honest, whereas a lack of reference to norms is not helpful in predicting veracity. Similarly, equivocation had a ‘hit’ rate of 81% for honest appeals, but only 56% for deceptive appeals, suggesting that appeals which contain large amounts of equivocal language are very likely to be deceptive, whereas a lack of equivocation is not helpful in predicting veracity.

To summarise, the present study could be considered to add to the emergent body of literature on real life, high stake deception, and have some potentially important implications. The findings suggest that there may be some verbal behaviours (the use of equivocal language, and speech errors) that have a general utility in discriminating between honesty and deception across high stakes situations. Moreover, some potentially context-specific cues were identified that, to the author’s knowledge, have not previously been investigated (expressions of hope of finding the missing relative alive, references to norms of behaviour and emotion, expressions of concern, pain, and grief, avoiding brutal language or detail, and positive emotion expressed towards the relative). These findings obviously require replication, and it is also unclear as to the extent to which the cues identified here are actually specific to the context of making an appeal for help with a missing or murdered relative. Nevertheless, there are perhaps indications here that a methodology utilising a tightly-focused context and a data-driven approach, may be useful in uncovering

behaviours that are helpful in discriminating between deception and honesty within specific contexts.

# **Chapter 10**

## **Study 6: Frequency Counts of Nonverbal Cues to Deception in Appeals, and Case by Case Analyses using Subjective Cues, Verbal Cues, and Nonverbal Cues**

### **10.1. Introduction**

In the introductory chapters of the present thesis, attention was drawn to the possible importance of considering multiple channels of behaviour when attempting to detect deception. Several leading researchers have suggested that the more cues to deception that a person exhibits, the more likely that person is to be lying (for example, DePaulo et al., 2003; ten Brinke & Porter, 2012; Vrij & Mann, 2004). The studies presented so far have examined subjective observer ratings, and verbal cues to deception; the study presented in this chapter, Study 6, adds to these by investigating nonverbal cues to deception in the same, larger sample of appeals used in Studies 4 and 5. As with verbal cues, some nonverbal behaviours are amenable to frequency counts and, therefore, may be more objectively measured, making them particularly useful for applied purposes.

As has previously been noted, the findings regarding several nonverbal cues to deception in high stakes contexts have been mixed; for example, gaze aversion has been found to relate to deception in some studies (Vrij & Mann, 2001a; Wright Whelan, 2009), but to have no relationship to veracity in another study (Mann et al., 2002). Similarly, the findings regarding hand movements and blink rate have been equally inconsistent: self-adaptors and illustrators were found to decrease during deception in one study (Koper & Sahlman, 1991), gesticulation was found to increase during deception in another study (Davis et al., 2005), and hand movements were found to be unaffected by veracity in another study (Mann et al., 2002); blink rate has been found to decrease (Mann et al., 2002), and to increase (ten Brinke & Porter, 2012), during high stakes deception. Head shaking has been found to relate to deception in two investigations of high stakes deceptive communication (Davis et al., 2005; Mann et al., 2006), and shrugging was related to deception in one study as part of a ‘nonverbal overdone’ measure (Davis et al., 2005). In general, the findings regarding nonverbal behaviours in high stakes contexts are scant and patchy, and there is clearly a need for further investigation of this channel of behaviour.

Nevertheless, on the basis of previous research, it was hypothesised that some nonverbal cues to deception, previously identified in high stakes situations, would generalise to this appeals context. Gaze aversion and head shaking, both of which have been shown to relate to deception across high stakes contexts, were particularly expected to relate to deception in the context of appeals. The inclusion of shrugging as a cue was more speculative, as there is very little previous research on this behaviour in high stakes contexts. The present study did not include an investigation of cues specifically derived from Study 1 of the present thesis, as the only nonverbal

cue amenable to frequency counts identified in Study 1 was gaze aversion, which, as mentioned above, has been identified in previous deception research.

As in Studies 2, 3, 4, and 5, case by case analyses were again conducted as it was hypothesised that the use of a combination of cues to classify cases (appeals) as honest or deceptive, would achieve high accuracy rates in discriminating between liars and truth-tellers. These analyses included combining verbal and nonverbal cues to classify cases, as it was considered that combining information from cues across two channels of behaviour, which are also amenable to frequency counts, would have the greatest utility in an applied context, and would be more accurate than case by case analyses using a single channel of behaviour. The final part of the case by case analyses involved combining the information from verbal cues, nonverbal cues, and cues subjectively rated by observers, to classify individual cases as honest or deceptive.

## **10.2 Method**

### **10.2.1. *Materials and Procedure***

The same 32 appeals for help with missing or murdered relatives (16 honest and 16 deceptive) used in Studies 4 and 5 were used in this study.

Again, following protocols established in previous research on high stakes deception (Mann et al., 2002; ten Brinke & Porter, 2012), one coder counted frequencies of predetermined behaviours in the video clips, and a second coder, who was blind to the veracity of the appeals, coded a random sample of 25% of the

deceptive appeals and 25% of the honest appeals. Both coders were graduate students studying deception. Inter-rater reliability for the frequency counts was high, ranging from  $r = .93$  to  $r = .1.00$  (see Table 10.1). As the inter-rater reliability for the frequency counts was high for the random sample, frequency counts produced by the coder who had coded all the transcripts were used in all analyses.

The behaviours included in the study were nonverbal cues to deception amenable to measurement by frequency counts, identified in previous research on real life, high stakes lies. The cues under consideration were in some ways proscribed by the footage available; for example, despite previous findings suggesting that blink rate may be related to deception in high stakes contexts (Mann et al., 2002; ten Brinke & Porter, 2012), in most cases the quality of the video clips was not of a sufficiently high standard to measure blink rate. Similarly, despite previous findings suggesting that the use of hand movements may be related to deception in high stakes contexts (Davis et al., 2005; Koper & Sahlman, 1991), in a large proportion of the appeals the hands were not visible, so that it was not possible to investigate the use of hand movements.

Gaze aversion was included as a cue in the present study, as the findings regarding this cue in the literature on high stakes deception have been mixed, and it therefore clearly warrants further investigation. Hence although in Study 4 of the present thesis, gaze aversion as a cue rated subjectively by observers did not achieve significant inter-rater reliability, it was considered useful to investigate this behaviour with more objective frequency counts. Appealers were coded as averting their gaze when they were not looking directly at the cameras or interviewer, but instead looked down, up or to the side. Appealers who were wearing sunglasses or hats so that their eyes were not visible, or whose eye-line was unclear, were not



included in the analysis (three deceptive appealers and two honest appealers were excluded for this reason). Appealers who were reading a statement were also not included in the analysis of gaze aversion, as this was considered to be a form of gaze aversion not indicative of factors that may be related to deception (for example, cognitive load or affective factors). One of the 16 deceptive appealers was reading a statement, and eight of the honest appealers were reading a statement.

Head shaking was included as a cue to deception, as increased headshaking was found to relate to deception in a study of police suspect interviews by Mann et al. (2006), and also in a study investigating cues to deception in criminal suspects' statements (Davis et al., 2005). In the same study, increased shrugging (as part of a group of 'nonverbal overdone' cues) was found to be marginally significantly related to deception, and so was also included as a cue in the present study. It was not possible to code shrugging in three of the deceptive appeals, as some of the footage was tight to the face of the appealers and so the shoulders were not visible. For a full description of each cue, and instructions to coders for frequency counts, see Appendix 10.

### **10.3. Results**

#### **10.3.1. *Analyses of Cues***

Preliminary analyses revealed no effect of the length of the appeals on veracity,  $F(1,31) = 0.10$ ,  $p = .757$ , and so this was not included as a variable. However, to account for the differing lengths of the appeals, each cue score for each appeal was

divided by the length (in seconds) of the appeal, to produce a frequency count for each cue as a proportion of the total appeal. For some appeals, the appealer was not visible for the entire length of the appeal as, for example, still shots of the victim were inserted in to the broadcast. In these cases, the cue scores were divided by the length of time (in seconds) for which the appealer was visible throughout the appeal. As this produced very small numbers, for the purposes of clarity this proportion was then multiplied by 100, which produced a frequency count per 100 seconds.

To investigate which cues were associated with deception, a series of one way ANOVAs, with veracity as a between-subjects independent variable, were conducted on the adjusted frequency counts on the cues. It was not possible to conduct a MANOVA, as not all cues were visible in all the appeals (it was not possible to measure gaze aversion and shrugging in every appeal, as discussed previously). Means, standard deviations, confidence intervals,  $F$  ratios,  $p$  values and effect sizes are presented in Table 10.1.

Analyses yielded significant effects for the cues gaze aversion and headshaking, both with means in the expected direction, i.e. deceptive appealers produced more gaze aversion and more head shaking than honest appealers. No effect of veracity was found for shrugging.

Levene's test for homogeneity of variance was significant for the cue headshaking. However, Welch's Test produced comparable results ( $p = .049$ ).

Due to the presence of extreme outliers on one of the variables, shrugging, notwithstanding arguments about the robustness of ANOVA when sample sizes are equal (Field, 2009), scores on this variable were also coded dichotomously as present/absent and a Fisher's Exact test was conducted on this cue. The analysis did

not yield a significant result ( $p > .05$ ), in line with the parametric analysis of this cue.

Table 10.1: *Inter-rater reliabilities, means (SDs), CIs, F ratios and p values, and effect sizes for frequency counts of nonverbal cues as a proportion of the total appeals*

Cue	Inter-rater reliability ( $r$ )	Deceptive Mean (SD) CI	Honest Mean (SD) CI	F ratio ( $p$ )	Effect size
Gaze aversion	$r = 1.00$ $p < .001$	<b>59.35</b> (30.50) 42.20 – 76.50	<b>23.53</b> (23.26) 1.07 – 45.98	$F(1,19) = 7.16$ $p = .016$	$d = 1.32$
Headshaking*	$r = .98$ $p < .001$	<b>25.26</b> (24.13) 15.37 – 35.14	<b>11.04</b> (12.94) 1.16 – 20.92	$F(1,31) = 4.32$ $p = .046$	$d = 0.73$
Shrugging	$r = .93$ $p = .001$	2.19 (4.33) -0.36 – 4.73	1.77 (4.58) -0.52 – 4.07	$F(1,29) = 0.06$ $p = .807$	$d = 0.09$

Means in bold differ significantly from each other, \* Levene's significant

### 10.3.2. Case by Case Analyses

#### 10.3.2.1. Verbal and nonverbal cues combined

As in Studies 2, 3, 4 and 5 of the present thesis, and in accordance with previous studies (Ekman et al., 1991; Leal et al., 2010), case by case analyses were conducted to assess the potential of each cue for lie detection. As both gaze aversion and head shaking were related to deception, cue scores for these two cues were classified as a 'hit' if they fell above the cue mean in deceptive appeals, and if they fell below the cue mean in honest appeals. In this way, a percentage of 'hits' was calculated for each cue (i.e. cases combined). Gaze aversion had a 'hit' rate of 79%, and head shaking had a 'hit' rate of 69%; see Table 10.2.

As only two nonverbal cues were found to discriminate between honest and deceptive appeals, and one of these cues (gaze aversion) was not available in all cases, the nonverbal cues were not considered separately to classify individual cases. Instead, they were combined with the verbal cues from the case by case analyses in Study 5. The cues considered in these analyses, then, were the verbal and nonverbal cues that had significantly discriminated between liars and truth-tellers on a two-tailed test, in the larger sample of appeals. These cues were: gaze aversion, head shaking, and the use of equivocal language (all related to deception); positive emotion expressed towards the relative, references to norms of emotion/behaviour, and, in cases where the relative was not publicly known to be dead, expressions of hope of finding the missing relative alive (all related to honesty). As before, for deceptive appeals, cue scores that fell above the cue mean on cues to deception, and below the cue mean on cues to honesty, were classified as a 'hit'. Similarly, for honest appeals, cue scores that fell above the cue mean on cues to honesty, and below the cue mean on cues to deception, were classified as a 'hit'. The cue 'expressions of positive emotion towards the relative' (which was coded dichotomously as present/absent), was classified as a 'hit' if present in an honest appeal, and a 'hit' if absent in a deceptive appeal.

The percentage of 'hits' was then calculated for each case (i.e. cues combined). Honest appeals had 'hit' rates between 20 - 100% (mean 70%), deceptive appeals had 'hit' rates between 17 - 100% (mean 70%), and the mean hit rate was 70% (see Table 10.2).

The information from the verbal and nonverbal cues was then combined to classify individual cases (appeals) as deceptive or honest. Cases which scored above 50% 'hits' on the cues were categorised as correctly classified (the appeal scored in

the expected direction on the majority of the cues). Cases which scored 50% ‘hits’ on the cues were categorised as unclassified, and cases that scored below 50% ‘hits’ on the cues were categorised as misclassified (the appeal did not score in the expected direction on the majority of the cues). In this way, 12 deceptive appeals were correctly classified (75%), 2 were unclassified (13%) and 2 were misclassified (13%). Of the honest appeals, 13 were correctly classified (81%), 1 was unclassified (6%) and 2 were misclassified (13%). Overall, therefore, using these criteria, 78% of cases were correctly classified, 9% were unclassified and 13% were misclassified; see Table 10.2.

Table 10.2: *Case by case analyses using verbal and nonverbal cues: score per case (appeal), and overall ‘hit’ rate. For each case: overall ‘hit’ rate of cues. All ‘hit’s related to overall cue means*

	Norms ( <i>M</i> = 0.53)	Positive emotion; dichotomo us	Hope ( <i>M</i> = 1.45)	Equivocal language ( <i>M</i> = 3.36)	Gaze aversion ( <i>M</i> = 46.15)	Head shaking ( <i>M</i> = 17.77)	Total hits
<b><u>Deceptive appeals</u></b>							
Case 10	<b>0.00</b>	<b>Absent</b>	1.88	<b>5.26</b>	Not visible	<b>43.93</b>	80%
Case 11	<b>0.00</b>	<b>Absent</b>	n/a	0.00	<b>72.73</b>	<b>22.73</b>	80%
Case 12	<b>0.00</b>	<b>Absent</b>	n/a	0.97	<b>77.94</b>	<b>47.06</b>	80%
Case 13	<b>0.00</b>	Present	<b>0.00</b>	<b>12.00</b>	Not visible	0.00	60%
Case 14	<b>0.00</b>	Present	n/a	<b>4.17</b>	<b>52.06</b>	<b>21.92</b>	80%
Case 15	<b>0.00</b>	<b>Absent</b>	n/a	1.03	Not visible	0.00	50%
Case 16	<b>0.00</b>	<b>Absent</b>	n/a	<b>7.41</b>	0.00	8.00	60%
Case 17	<b>0.00</b>	<b>Absent</b>	<b>0.00</b>	<b>14.29</b>	20.83	<b>75.00</b>	83%
Case 18	<b>0.00</b>	Present	<b>0.00</b>	0.00	Not visible	0.00	40%
Case 19	<b>0.00</b>	<b>Absent</b>	<b>0.00</b>	<b>6.90</b>	<b>82.86</b>	<b>22.86</b>	100%
Case 20	<b>0.00</b>	<b>Absent</b>	<b>0.00</b>	0.00	<b>70.69</b>	<b>20.69</b>	83%
Case 21	<b>0.00</b>	<b>Absent</b>	2.44	<b>4.88</b>	<b>66.67</b>	0.00	67%

Case 22	<b>0.00</b>	Present	<b>0.00</b>	1.87	<b>100.00</b>	17.69	50%
Case 23	0.91	Present	<b>0.46</b>	1.37	23.79	6.80	17%
Case 24	<b>0.00</b>	<b>Absent</b>	1.49	<b>8.96</b>	<b>90.53</b>	<b>67.37</b>	83%
Case 25	<b>0.30</b>	<b>Absent</b>	n/a	<b>5.95</b>	<b>54.13</b>	<b>44.63</b>	100%
<b>Cue hit rate: decept. appeals</b>	94%	69%	70%	56%	75%	56%	
	Norms ( <i>M</i> = 0.53)	Positive emotion; dichotomous	Hope ( <i>M</i> = 1.45)	Equivocal language ( <i>M</i> = 3.36)	Gaze aversion ( <i>M</i> = 46.15)	Head shaking ( <i>M</i> = 17.77)	Total hits
<b><u>Honest appeals</u></b>							
Case 26	0.00	Absent	<b>2.11</b>	<b>2.11</b>	Not visible	<b>0.00</b>	60%
Case 27	0.00	Absent	<b>4.76</b>	7.94	<b>43.48</b>	<b>5.80</b>	50%
Case 28	<b>2.53</b>	<b>Present</b>	1.27	<b>2.53</b>	<b>0.00</b>	33.90	67%
Case 29	<b>0.70</b>	<b>Present</b>	n/a	<b>0.00</b>	Not visible	<b>11.72</b>	100%
Case 30	<b>3.23</b>	Absent	n/a	<b>1.61</b>	Not visible	<b>0.00</b>	75%
Case 31	0.00	<b>Present</b>	1.18	3.53	Not visible	30.44	20%
Case 32	<b>0.83</b>	<b>Present</b>	n/a	<b>1.65</b>	<b>26.67</b>	<b>8.89</b>	100%
Case 33	0.00	<b>Present</b>	n/a	<b>3.03</b>	Not visible	<b>8.33</b>	75%
Case 34	<b>5.56</b>	Absent	n/a	4.17	63.22	36.78	20%
Case 35	0.00	Absent	<b>2.27</b>	<b>0.00</b>	Not visible	<b>5.13</b>	60%
Case 36	<b>1.36</b>	<b>Present</b>	n/a	<b>1.36</b>	Not visible	<b>0.00</b>	100%
Case 37	0.00	<b>Present</b>	0.00	<b>0.00</b>	Not visible	<b>12.70</b>	60%
Case 38	0.00	<b>Present</b>	0.00	<b>2.33</b>	<b>19.67</b>	<b>16.39</b>	67%
Case 39	0.00	<b>Present</b>	<b>3.61</b>	<b>0.00</b>	<b>11.65</b>	<b>0.00</b>	83%
Case 40	<b>1.43</b>	<b>Present</b>	<b>5.71</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	100%
Case 41	0.00	<b>Present</b>	<b>1.76</b>	<b>2.20</b>	Not visible	<b>0.00</b>	83%
<b>Cue hit rate: honest appeals</b>	44%	69%	60%	81%	86%	81%	
<b>TOTAL HITS:</b>	69%	69%	65%	69%	79%	69%	

Scores in **bold** are 'hits', i.e. case score is in expected direction.

#### 10.3.2.2. *Verbal, nonverbal, and subjective cues combined*

The next part of the case by case analyses was to combine the information from the verbal cues, nonverbal cues, and cues subjectively rated by observers. The same procedure was used as previously to classify individual cases as deceptive or honest. Cues included were those that had significantly discriminated between honest and deceptive appeals on a two-tailed test. Cues related to deception were: gaze aversion, head shaking, the use of equivocal language, fake emotion, fake facial expression, putting on a performance, dislike for the appealer, creepy, behaving unnaturally, and no sympathy for the appealer. Cues related to honesty were: positive emotion expressed towards the relative, references to norms of emotion/behaviour, expressions of hope, genuinely sad, sad eyes, genuine and heartfelt appeal, feeling the appealer's pain, feeling sorry for the appealer, normal, urgency, personal and expressive voice, plausible, and voice quivering with genuine emotion.

Honest appeals had 'hit' rates between 23 - 100% (mean 73%), deceptive appeals had 'hit' rates between 17 - 100% (mean 66%), and the mean 'hit' rate was 70% (see Table 10.3). Again, cases which scored above 50% 'hits' on the cues were categorised as correctly classified, cases which scored 50% 'hits' on the cues were categorised as unclassified, and cases that scored below 50% 'hits' on the cues were categorised as misclassified. In this way, 10 deceptive appeals were correctly classified (63%), 0 were unclassified, and 6 were misclassified (38%). Of the honest appeals, 12 were correctly classified (75%), 1 was unclassified (6%) and 3 were misclassified (19%). Overall, therefore, using these criteria, 22 cases (69%) were correctly classified, 1 (3%) was unclassified and 9 (28%) were misclassified.

Table 10.3: Case by case analyses; verbal cues, nonverbal cues, and cues rated by observers. Overall cue mean, score per case (appeal), and overall 'hit' rate. For each case: 'hit' rate of cues. All 'hit's related to overall cue means.

	Nor ms	Hope	Positiv e/ love	Equi vocat ion	Gaze avers ion	Head shake	Fake emot ion	Fake facial expre ssion	Perfo rman ce	Unna tural	Cree py	Disli ke	No symp athy	Sad eyes	Genu inely sad	Genu ine / heart felt	Urge ncy	Plaus ible	Nor mal	Perso nal & expre ssive	Voca l quive r	Feel pain	Feel sorry	Total hits
	<i>M</i> = 0.53	<i>M</i> = 1.45	dichoto mous	<i>M</i> = 3.36	<i>M</i> = 46.15	<i>M</i> = 17.77	<i>M</i> = 2.91	<i>M</i> = 2.92	<i>M</i> = 2.99	<i>M</i> = 3.08	<i>M</i> = 2.95	<i>M</i> = 3.08	<i>M</i> = 2.91	<i>M</i> = 2.94	<i>M</i> = 3.08	<i>M</i> = 2.98	<i>M</i> = 3.33	<i>M</i> = 3.01	<i>M</i> = 3.04	<i>M</i> = 3.08	<i>M</i> = 2.96	<i>M</i> = 3.03	<i>M</i> = 3.08	
<b>Deceptive Appeals</b>																								
Case 10	<b>0.00</b>	1.88	<b>absent</b>	<b>5.26</b>	Not vis.	<b>43.93</b>	<b>4.00</b>	<b>3.60</b>	<b>3.80</b>	<b>3.60</b>	2.60	<b>3.40</b>	<b>3.60</b>	<b>2.20</b>	<b>2.00</b>	<b>1.80</b>	<b>2.20</b>	<b>2.60</b>	<b>2.60</b>	<b>2.20</b>	<b>1.40</b>	<b>2.40</b>	<b>2.20</b>	91%
Case 11	<b>0.00</b>	n/a	<b>absent</b>	0.00	<b>72.73</b>	<b>22.73</b>	<b>4.20</b>	<b>4.00</b>	<b>4.20</b>	<b>4.40</b>	<b>4.40</b>	<b>4.00</b>	<b>4.00</b>	<b>1.20</b>	<b>1.40</b>	<b>1.20</b>	<b>3.20</b>	<b>1.60</b>	<b>2.20</b>	<b>1.80</b>	<b>1.40</b>	<b>1.60</b>	<b>2.20</b>	95%
Case 12	<b>0.00</b>	n/a	<b>absent</b>	0.97	<b>77.94</b>	<b>47.6</b>	<b>4.40</b>	<b>4.40</b>	<b>4.60</b>	<b>4.80</b>	<b>3.80</b>	<b>4.20</b>	<b>4.00</b>	<b>1.40</b>	<b>1.40</b>	<b>1.40</b>	<b>1.60</b>	<b>2.00</b>	<b>2.00</b>	<b>2.40</b>	<b>1.40</b>	<b>1.60</b>	<b>1.40</b>	95%
Case 13	<b>0.00</b>	<b>0.00</b>	present	<b>12.00</b>	Not vis.	0.00	2.80	2.80	<b>3.00</b>	2.60	<b>3.20</b>	3.00	2.80	3.00	3.40	3.60	<b>3.00</b>	3.40	<b>3.00</b>	<b>3.00</b>	<b>2.60</b>	3.20	3.60	41%
Case 14	<b>0.00</b>	n/a	present	<b>4.17</b>	<b>52.06</b>	<b>21.92</b>	2.80	2.80	2.60	<b>3.60</b>	<b>3.00</b>	<b>3.20</b>	<b>3.00</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	<b>2.80</b>	3.40	3.60	<b>2.60</b>	<b>2.40</b>	<b>2.80</b>	<b>2.80</b>	73%
Case 15	<b>0.00</b>	n/a	<b>absent</b>	1.03	Not vis.	0.00	<b>3.20</b>	<b>3.00</b>	<b>3.20</b>	<b>3.40</b>	<b>3.60</b>	<b>3.20</b>	<b>3.20</b>	3.00	<b>2.80</b>	<b>2.60</b>	<b>2.60</b>	<b>2.00</b>	<b>2.40</b>	<b>2.40</b>	<b>2.60</b>	<b>2.80</b>	<b>3.00</b>	86%
Case 16	<b>0.00</b>	n/a	<b>absent</b>	<b>7.41</b>	0.00	8.00	2.60	2.60	2.80	3.00	<b>4.20</b>	3.00	2.60	4.00	3.60	3.40	4.40	3.20	<b>1.80</b>	3.60	3.60	3.40	3.60	23%
Case 17	<b>0.00</b>	<b>0.00</b>	<b>absent</b>	<b>14.29</b>	20.83	<b>75.00</b>	<b>4.33</b>	<b>4.33</b>	<b>4.00</b>	<b>4.00</b>	<b>3.33</b>	<b>4.33</b>	<b>4.00</b>	<b>1.00</b>	<b>1.67</b>	<b>1.67</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>1.67</b>	<b>1.67</b>	<b>1.67</b>	96%
Case 18	<b>0.00</b>	<b>0.00</b>	present	0.00	Not vis.	0.00	<b>3.60</b>	<b>3.80</b>	<b>3.80</b>	<b>3.20</b>	<b>3.20</b>	<b>3.40</b>	<b>3.60</b>	<b>1.40</b>	<b>2.60</b>	<b>2.60</b>	3.40	<b>2.20</b>	<b>2.80</b>	<b>2.80</b>	<b>2.60</b>	<b>2.40</b>	<b>2.60</b>	82%
Case 19	<b>0.00</b>	<b>0.00</b>	<b>absent</b>	<b>6.90</b>	<b>82.86</b>	<b>22.86</b>	2.80	<b>3.20</b>	<b>3.00</b>	3.00	2.20	<b>3.20</b>	<b>3.00</b>	3.40	3.20	3.00	<b>2.60</b>	3.20	3.20	3.20	3.40	3.20	3.20	48%
Case 20	<b>0.00</b>	<b>0.00</b>	<b>absent</b>	0.00	<b>70.69</b>	<b>20.69</b>	2.60	2.40	2.80	2.80	2.80	2.60	2.60	3.60	3.60	3.40	4.20	3.60	3.60	3.40	3.20	3.60	3.40	22%
Case 21	<b>0.00</b>	2.44	<b>absent</b>	<b>4.88</b>	<b>66.67</b>	0.00	2.60	2.60	2.40	3.00	2.60	2.80	2.60	3.60	3.40	3.60	3.40	3.40	3.80	3.40	3.00	3.60	3.40	17%
Case 22	<b>0.00</b>	<b>0.00</b>	present	1.87	<b>100.0</b>	17.69	1.80	2.20	2.20	2.00	2.20	2.20	1.80	4.40	4.20	4.00	4.40	3.60	3.40	4.20	4.00	4.00	3.80	17%



	Nor ms	Hope	Positiv e/ love	Equi vocat ion	Gaze avers ion	Head shake	Fake emot ion	Fake facial expre ssion	Perfo rman ce	Unna tural	Cree py	Disli ke	No symp athy	Sad eyes	Genu inely sad	Genu ine / heart felt	Urge ncy	Plaus ible	Nor mal	Perso nal & expre ssive	Voca l quive r	Feel pain	Feel sorry	Total hits
	<i>M</i> = 0.53	<i>M</i> = 1.45	dichoto mous	<i>M</i> = 3.36	<i>M</i> = 46.15	<i>M</i> = 17.77	<i>M</i> = 2.91	<i>M</i> = 2.92	<i>M</i> = 2.99	<i>M</i> = 3.08	<i>M</i> = 2.95	<i>M</i> = 3.08	<i>M</i> = 2.91	<i>M</i> = 2.94	<i>M</i> = 3.08	<i>M</i> = 2.98	<i>M</i> = 3.33	<i>M</i> = 3.01	<i>M</i> = 3.04	<i>M</i> = 3.08	<i>M</i> = 2.96	<i>M</i> = 3.03	<i>M</i> = 3.08	
Case 23	0.91	<b>0.46</b>	present	1.37	23.76	6.80	<b>4.20</b>	<b>4.20</b>	<b>4.00</b>	<b>4.00</b>	<b>3.80</b>	<b>4.40</b>	<b>3.60</b>	<b>1.80</b>	<b>2.20</b>	<b>1.80</b>	<b>3.20</b>	<b>2.20</b>	<b>2.80</b>	<b>2.40</b>	<b>1.80</b>	<b>2.20</b>	<b>2.20</b>	78%
Case 24	<b>0.00</b>	1.49	absent	<b>8.96</b>	<b>90.53</b>	<b>67.37</b>	<b>3.60</b>	<b>3.80</b>	<b>3.80</b>	<b>3.60</b>	<b>3.40</b>	<b>3.80</b>	<b>3.40</b>	<b>2.20</b>	<b>2.20</b>	<b>2.00</b>	<b>2.60</b>	<b>2.60</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	96%
Case 25	<b>0.30</b>	n/a	absent	<b>5.95</b>	<b>54.13</b>	<b>44.63</b>	<b>3.80</b>	<b>3.80</b>	<b>3.80</b>	<b>3.40</b>	<b>4.40</b>	<b>4.00</b>	<b>4.00</b>	<b>2.20</b>	<b>2.20</b>	<b>2.40</b>	<b>2.60</b>	<b>2.40</b>	<b>2.20</b>	<b>2.20</b>	<b>2.60</b>	<b>2.20</b>	<b>1.80</b>	100 %
Hits decep. Appeals	94%	70%	69%	56%	75%	56%	56%	63%	69%	63%	69%	69%	69%	56%	63%	63%	69%	56%	69%	69%	69%	63%	63%	

#### Honest Appeals

Case 26	0.00	<b>2.11</b>	absent	<b>2.11</b>	Not vis.	<b>0.00</b>	3.00	3.40	3.00	3.20	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	2.60	3.00	<b>3.00</b>	2.60	<b>3.20</b>	<b>3.60</b>	<b>3.20</b>	2.20	<b>3.20</b>	<b>3.40</b>	55%
Case 27	0.00	<b>4.76</b>	absent	7.94	<b>43.48</b>	<b>5.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.60</b>	<b>2.80</b>	<b>2.60</b>	<b>3.20</b>	2.80	<b>3.00</b>	2.60	<b>3.20</b>	<b>3.60</b>	2.80	2.60	2.80	3.00	61%
Case 28	<b>2.53</b>	1.27	<b>presen t</b>	<b>2.53</b>	<b>0.00</b>	33.90	<b>2.60</b>	<b>2.80</b>	<b>2.80</b>	<b>2.80</b>	<b>2.60</b>	<b>2.60</b>	<b>2.60</b>	2.60	<b>3.40</b>	<b>3.00</b>	<b>3.40</b>	<b>3.60</b>	<b>3.60</b>	<b>3.80</b>	<b>3.20</b>	3.00	<b>3.40</b>	83%
Case 29	<b>0.70</b>	n/a	<b>presen t</b>	<b>0.00</b>	Not vis.	<b>11.72</b>	<b>1.80</b>	<b>1.40</b>	<b>2.20</b>	<b>2.20</b>	<b>2.20</b>	<b>2.40</b>	<b>2.00</b>	<b>4.20</b>	<b>4.20</b>	<b>4.40</b>	<b>4.60</b>	<b>3.40</b>	<b>3.60</b>	<b>4.40</b>	<b>4.40</b>	<b>3.80</b>	<b>4.20</b>	100 %
Case 30	<b>3.23</b>	n/a	absent	<b>1.61</b>	Not vis.	<b>0.00</b>	<b>2.80</b>	<b>2.60</b>	<b>2.60</b>	3.40	3.00	<b>2.60</b>	<b>2.80</b>	<b>3.40</b>	<b>3.40</b>	<b>3.40</b>	<b>3.40</b>	<b>3.20</b>	<b>3.20</b>	<b>3.20</b>	<b>3.80</b>	<b>3.40</b>	<b>3.80</b>	86%
Case 31	0.00	1.18	<b>presen t</b>	3.53	Not vis.	30.44	3.00	3.00	3.20	3.20	<b>2.40</b>	<b>3.00</b>	3.00	<b>3.00</b>	2.60	2.60	3.20	3.00	<b>3.40</b>	3.00	2.40	2.80	3.00	23%
Case 32	<b>0.83</b>	n/a	<b>presen t</b>	<b>1.65</b>	<b>26.67</b>	<b>8.89</b>	<b>1.80</b>	<b>1.80</b>	<b>2.00</b>	<b>2.40</b>	<b>2.20</b>	<b>1.80</b>	<b>2.00</b>	<b>4.00</b>	<b>4.40</b>	<b>4.40</b>	<b>4.80</b>	<b>4.00</b>	<b>3.80</b>	<b>4.00</b>	<b>4.00</b>	<b>4.20</b>	<b>4.00</b>	100 %
Case 33	0.00	n/a	<b>presen t</b>	<b>3.03</b>	Not vis.	<b>8.33</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>2.00</b>	<b>2.40</b>	<b>2.20</b>	<b>2.00</b>	<b>3.20</b>	<b>3.80</b>	<b>3.80</b>	<b>3.40</b>	<b>3.80</b>	<b>3.80</b>	<b>3.80</b>	<b>4.00</b>	<b>4.00</b>	<b>3.80</b>	95%
Case 34	<b>5.56</b>	n/a	absent	4.17	63.22	36.78	3.00	3.00	3.00	3.40	3.40	3.40	<b>2.80</b>	<b>3.00</b>	<b>3.20</b>	<b>3.00</b>	<b>4.60</b>	2.80	2.40	3.00	<b>3.40</b>	<b>3.20</b>	3.00	36%

	Nor ms	Hope	Positiv e/ love	Equi vocat ion	Gaze avers ion	Head shake	Fake emot ion	Fake facial expre ssion	Perfo rman ce	Unna tural	Cree py	Disli ke	No symp athy	Sad eyes	Genu inely sad	Genu ine / heart felt	Urge ncy	Plaus ible	Nor mal	Perso nal & expre ssive	Voca l quive r	Feel pain	Feel sorry	Total hits
	<i>M</i> = 0.53	<i>M</i> = 1.45	dichoto mous	<i>M</i> = 3.36	<i>M</i> = 46.15	<i>M</i> = 17.77	<i>M</i> = 2.91	<i>M</i> = 2.92	<i>M</i> = 2.99	<i>M</i> = 3.08	<i>M</i> = 2.95	<i>M</i> = 3.08	<i>M</i> = 2.91	<i>M</i> = 2.94	<i>M</i> = 3.08	<i>M</i> = 2.98	<i>M</i> = 3.33	<i>M</i> = 3.01	<i>M</i> = 3.04	<i>M</i> = 3.08	<i>M</i> = 2.96	<i>M</i> = 3.03	<i>M</i> = 3.08	
Case 35	0.00	<b>2.27</b>	absent	<b>0.00</b>	Not vis.	<b>5.13</b>	<b>2.40</b>	<b>2.40</b>	<b>2.40</b>	3.80	<b>2.80</b>	3.40	<b>2.80</b>	<b>3.20</b>	3.00	2.60	3.00	<b>3.40</b>	3.00	2.40	2.60	2.80	<b>3.20</b>	50%
Case 36	<b>1.36</b>	n/a	<b>presen t</b>	<b>1.36</b>	Not vis.	<b>0.00</b>	<b>2.80</b>	<b>2.80</b>	3.40	<b>2.80</b>	<b>2.20</b>	<b>3.00</b>	<b>2.80</b>	2.80	<b>3.40</b>	<b>3.20</b>	2.60	3.00	<b>3.40</b>	2.80	<b>3.00</b>	2.60	<b>3.20</b>	71%
Case 37	0.00	0.00	<b>presen t</b>	<b>0.00</b>	Not vis.	<b>12.70</b>	<b>2.20</b>	<b>2.00</b>	<b>2.00</b>	<b>1.80</b>	<b>2.40</b>	<b>2.40</b>	<b>2.20</b>	<b>4.20</b>	<b>4.20</b>	<b>4.00</b>	<b>4.60</b>	<b>3.60</b>	<b>3.20</b>	<b>4.00</b>	<b>4.20</b>	<b>3.80</b>	<b>3.80</b>	91%
Case 38	0.00	0.00	<b>presen t</b>	<b>2.33</b>	<b>19.67</b>	<b>16.39</b>	3.00	3.00	3.20	3.40	<b>2.80</b>	3.20	3.20	<b>3.00</b>	<b>3.20</b>	2.80	<b>3.60</b>	2.60	3.00	2.80	<b>3.00</b>	2.80	2.80	39%
Case 39	0.00	<b>3.61</b>	<b>presen t</b>	<b>0.00</b>	<b>11.65</b>	<b>0.00</b>	<b>2.40</b>	<b>2.60</b>	<b>2.80</b>	<b>2.20</b>	3.60	3.20	<b>2.60</b>	<b>3.60</b>	<b>3.80</b>	<b>3.60</b>	<b>4.00</b>	<b>3.40</b>	2.60	<b>3.40</b>	<b>3.60</b>	<b>3.40</b>	<b>3.20</b>	83%
Case 40	<b>1.43</b>	<b>5.71</b>	<b>presen t</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>1.60</b>	<b>1.40</b>	<b>1.60</b>	<b>1.80</b>	<b>1.80</b>	<b>2.20</b>	<b>2.00</b>	<b>4.00</b>	<b>4.40</b>	<b>4.40</b>	<b>4.40</b>	<b>3.40</b>	<b>4.00</b>	<b>4.60</b>	<b>4.60</b>	<b>4.60</b>	<b>4.60</b>	100 %
Case 41	0.00	<b>1.76</b>	<b>presen t</b>	<b>2.20</b>	Not vis.	<b>0.00</b>	<b>2.60</b>	<b>2.80</b>	<b>2.80</b>	<b>3.00</b>	<b>2.60</b>	<b>2.80</b>	3.00	<b>3.20</b>	<b>3.40</b>	<b>3.20</b>	<b>3.60</b>	<b>3.20</b>	<b>3.40</b>	<b>3.20</b>	<b>3.60</b>	<b>3.20</b>	<b>3.20</b>	91%
Hits hon. appeals	44%	60%	69%	81%	86%	81%	75%	75%	69%	63%	81%	75%	81%	81%	75%	81%	69%	75%	75%	63%	75%	63%	75%	
TOTAL HITS:	69%	65%	69%	69%	79%	69%	66%	69%	69%	63%	75%	72%	75%	69%	69%	72%	69%	66%	72%	66%	75%	63%	69%	

Scores in **bold** are ‘hits’, i.e. case score is in expected direction

### 10.3.2.3. Comparing the classification accuracy of the different groups of cues

The aim of the final part of the case by case analyses was to investigate whether using any particular combination of types of cue would be most effective for accurately classifying individual cases as honest or deceptive. So, for example, would using the subjective cues grouped as four factors (as investigated in Study 4) be significantly more or less effective for correctly classifying appeals than using verbal cues. A series of Chi-square comparisons ( $df = 2$ ) showed only one significant result; the verbal + nonverbal cues combination predicted more correct classifications than the subjective cues as four factors combination,  $\chi^2 = 6.93$ ,  $p = .031$ . None of the other comparisons was significant ( $p > .05$ ), though the trend was for the verbal + nonverbal combination to be very slightly superior. See Table 10.4.

Table 10.4: *Classifications (correctly classified, misclassified, and unclassified) of appeals using different groups of cues. Classifications shown as percentages, followed by number of cases classified in brackets*

	<b>Deceptive appeals (<i>n</i> = 16)</b>	<b>Honest appeals (<i>n</i> = 16)</b>	<b>All appeals (<i>n</i> = 32)</b>
<u>Subjective cues</u>			
Correctly classified	63% (10 cases)	75% (12 cases)	69% (22 cases)
Misclassified	37% (6 cases)	25% (4 cases)	31% (10 cases)
Unclassified	0% (0 cases)	0% (0 cases)	0% (0 cases)
<u>Subjective cues as four factors</u>			
Correctly classified	63% (10 cases)	63% (10 cases)	63% (20 cases)
Misclassified	37% (6 cases)	31% (5 cases)	34% (11 cases)
Unclassified	0% (0 cases)	6% (1 case)	3% (1 case)
<u>Verbal cues</u>			
Correctly classified	81% (13 cases)	56% (9 cases)	69% (22 cases)
Misclassified	6% (1 case)	19% (3 cases)	13% (4 cases)
Unclassified	13% (2 cases)	25% (4 cases)	19% (6 cases)
<u>Verbal + nonverbal cues</u>			

Correctly classified	75% (12 cases)	81% (13 cases)	78% (25 cases)
Misclassified	13% (2 cases)	13% (2 cases)	13% (4 cases)
Unclassified	13% (2 cases)	6% (1 case)	9% (3 cases)
<u>Verbal + nonverbal</u>			
<u>+ subjective cues</u>			
Correctly classified	63% (10 cases)	75% (12 cases)	69% (22 cases)
Misclassified	37% (6 cases)	19% (3 cases)	28% (9 cases)
Unclassified	0% (0 cases)	6% (1 case)	3% (1 case)

---

#### 10.4. Discussion

The findings of the present study could be considered to make a useful contribution to the small body of existing research which has investigated nonverbal cues to deception in high stakes situations.

The findings regarding gaze aversion are perhaps of particular interest, given the very substantial effect size, and that, in the present thesis, it was the most successful cue at predicting veracity (79% correct). As previously noted, despite the dominant pan-cultural stereotype that liars avert their gaze (The Global Deception Research Team, 2006), gaze aversion as a cue to deception has often been dismissed as a myth (for example, Vrij, 2004), as low stakes experimental findings do not support the belief that liars avert their gaze (DePaulo et al., 2003). Some researchers have even reported that liars may increase eye contact, in an attempt to appear credible, or monitor how their receiver is responding to their lies (Mann et al., 2012). Nevertheless, as noted earlier, there are strong theoretical grounds for the expectation

for gaze aversion to increase during deception, as gaze aversion has been related to several factors which are believed to underlie the production of cues to deception: gaze aversion has been variously related to cognitive load (Vrij & Mann, 2001a), emotional responses such as shame (DePaulo et al., 2003), and guilt (Ekman, 1992), and attempts by liars to be evasive and distance themselves from their communications (Zuckerman et al., 1981), and these factors may all be exacerbated in high stakes situations.

Meta-analytic findings suggest that gaze aversion emerges as a cue to deception when motivation to succeed in the lie is a moderating variable (DePaulo et al., 2003), and findings from previous research on high stakes deception suggest that, in some circumstances, liars do exhibit increased gaze aversion: for example, Vrij & Mann (2001a), in their analysis of the behaviour of a murder suspect, report that he displayed more gaze aversion when being deceptive in the section of the police interview before he confessed (i.e. when he still had high motivation to succeed in the lie). Also, Wright Whelan (2009) found that lying appealers were perceived as producing more gaze aversion than honest appealers. However, the findings regarding this behaviour have not been consistent; hence Mann et al. (2002), in their analysis of the behaviours of suspects in police interviews, found no significant effect of veracity on gaze aversion.

It could be the case, however, that gaze aversion is particularly prevalent in the context of appeals. For instance, when somebody is lying about killing his/her own relative, we might expect affective responses of shame and guilt to be

particularly prominent. Hence, in this particular context, increased gaze aversion may also be used as a predominant form of distancing behaviour. Honest appeals are likely to be a genuine attempt to communicate with the public and ask for help, hence honest appealers are more likely to behave in a way that facilitates this communication; for example, by using direct eye contact. The motivation of deceptive appealers in this context is clearly not the same, as they do not actually want the public to assist in finding out what happened to their relative. Deceptive appealers may, therefore, produce behaviours that inhibit direct communication, such as, gaze aversion.

The finding that deceptive appealers produce more head shaking than honest appealers is also in line with some previous research on real life, high stakes lies (Davis et al., 2005; Mann et al., 2006). Some researchers have suggested that this may be a piece of deliberately exaggerated behaviour intended to convey an appearance of honesty; i.e. an example of ‘protesting too much’ in a Shakespearean sense (Davis et al., 2005). Alternatively, or additionally, head shaking could be construed as a ‘leakage’ cue (Ekman, 1992); i.e. a manifestation of the negative emotions that the appealers are experiencing as a consequence of his or her deception. It is not clear from the present study which explanation may be more appropriate; the head shakes varied in magnitude from very small to quite large movements, and it could be that the former are more indicative of leakage, and the latter of exaggerated behaviour. This possible distinction may be worth following up in future research.

According to the case by case analyses, there was a trend indicating some advantage in using the combination of verbal and nonverbal cues for classifying individual cases as honest or deceptive, although this advantage over the other groups of cues/combinations of groups of cues was statistically significant only for subjective cues as four factors. Using this combination of verbal and nonverbal cues, of the 28 cases classified, 25 cases were correctly classified. This suggests that the approach of gathering information from multiple channels of behaviour, and using the kinds of cues identified here, may be worth investigating further. Notably, the only individual cue that was equivalent in terms of the ability to predict veracity was gaze aversion, but it was not possible to measure this in every case. This illustrates a general problem with using a restricted range of cues; i.e. it may not always be possible to identify them in practice. Given that researchers such as ten Brinke and Porter (2012), and Vrij and Mann (2004) have argued a multiple-cue approach is more likely to be more successful in the field, hopefully, the present research has possibly pointed to new directions to explore in this respect.

The combination of the subjective, verbal, and nonverbal cues was not significantly more accurate than any other combination of cues. However, interestingly, using the combination of verbal and nonverbal cues measured by frequency counts, two of the deceptive appeals were misclassified (cases 18 and 23), and two were unclassified (cases 15 and 22), and yet both the misclassified cases scored very highly in the expected directions on the subjective cues (94% and 100% respectively), as did one of the unclassified cases (94%). This suggests that there

may be elements of deceptive behaviour that, although not picked up by the verbal and nonverbal cues under consideration in the present thesis, are captured by more subjective observer impressions.

Given the variety of cues that were identified as of possible use in the case by case analyses, it is difficult to explain them in terms of any single, overarching, theoretical perspective, although some are more readily accommodated by existing theoretical viewpoints than others. For example, as previously noted, gaze aversion may be related to cognitive load or affective responses of shame, head shaking may be the result of leakage of negative emotion, and equivocation may be a form of psychological evasion from feelings of guilt. However, although to some extent it may be possible to explain the aforementioned cues in terms of general factors such as cognitive load, negative affective reaction, leakage, and distancing, in addition to these more familiar cues, the results of the studies presented thus far also draw attention to some more novel cues that are less obviously explicable in these ways, and might be more specifically related to the context of making an appeal. The theoretical implications of this are addressed in the general discussion.

We now turn our attention to the other side of the same coin; if it is the case the high stakes situation used here may produce some discernible cues to deception, how good are different samples of observers at recognising them?



# **Chapter 11**

## **Study 7: Police and Public Observer Accuracy in Detecting Deception in Appeals**

### **11.1. Introduction**

As has been mentioned on many occasions in this thesis, in general, people are poor at detecting deception, generally scoring barely above chance (see for example, Bond and DePaulo, 2006; Burgoon et al., 2008; O'Sullivan et al., 1988). However, according to the arguments put forward in this thesis, the main explanation for the apparent lack of observer accuracy is the fact that the large majority of studies of deception detection have used low stakes lies as stimulus materials, in which the motivation to lie is low, and the consequences of not being believed are minimal; hence, in typical laboratory studies, the stakes are not high enough to elicit discernible cues to deception, and thus deceit is almost impossible to detect (see, for example, Bond and DePaulo, 2006; Miller and Stiff, 1993).

Correspondingly, as noted in the introductory chapters, the few studies that have investigated observer accuracy in detecting deception in high stakes situations, have produced more promising results. For example, in a series of studies using videos of police interviews with suspects as stimulus materials (Mann & Vrij, 2006; Mann et al., 2004, 2006; Vrij & Mann, 2001b; Vrij et al., 2006), police officers' credibility judgements were consistently accurate at rates well above chance levels. These findings could be considered particularly important as previous studies using low stakes stimulus materials have suggested that police officers are poor at detecting deception (for example, DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991; Meissner & Kassin, 2002; Vrij & Graham, 1997). However, due to restrictions placed on the viewing of the videos of police interviews with suspects in the above series of studies, the authors were unable to test whether non-police participants would also achieve accuracy rates above chance levels in detecting deception. It is, therefore, not clear whether the high accuracy rates were due to the use of high stakes lies per se, or due to the cohort used as participants; it could be that police officers are particularly good at detecting this type of lie, perhaps because of their greater domain familiarity with the stimulus materials.

Of further interest in this small body of studies using high stakes stimulus materials and police officers as participants, are variations in the presence of truth bias; i.e. the tendency to judge a communication to be honest rather than deceptive, such that truthful messages are correctly identified more frequently than deceptive messages. As previously noted, in three of the studies no truth bias was found (Mann

& Vrij, 2006; Mann et al., 2004; Vrij et al., 2006); however, in two others some evidence of truth bias was found (Mann et al., 2006; Vrij & Mann, 2001a). To the author's knowledge, there have been no published investigations of deception detection accuracy and truth bias in the general public using high stakes lies as stimulus materials. Clearly, further investigation is needed into these issues.

Also, as noted previously, findings regarding the relationship between observer accuracy and confidence in credibility judgements in high stakes deception research have been mixed. In two studies no relationship was found between confidence and accuracy (Mann et al., 2004; Vrij & Mann, 2001a); however, in another study participants were more confident in their correct judgements than in their incorrect judgements of veracity (Mann et al., 2006). This finding in particular is interesting, as a review of the research in low stakes situations has suggested that there is no relationship between accuracy and confidence, so that when people attempt to detect deception, they report feeling as confident in their incorrect decisions, as in their correct decisions (DePaulo et al., 1997). Moreover, reports suggest that police officers tend to be more confident overall in their judgements than laypersons, despite being no more accurate (for example, DePaulo & Pfeifer, 1986; Meissner & Kassin, 2002). Again, to the author's knowledge, there have been no published investigations of the relationship between accuracy and confidence in the general public, using high stakes lies as stimulus materials. If observers are able to base their credibility judgements on more pronounced cues to deception, they may be more accurate and also more confident in their decisions.

Another issue that has not been addressed so far in any deception research, is the value of consensus judgements: if there is a strong consensus amongst observers that a 'sender' is lying, or telling the truth, how likely is it that the consensus judgement is correct? Most of the work in consensus effects psychology generally has concentrated on group influences on accuracy in decision making, in which members are required to make group decisions. The general finding is that group decisions where there is consensus are no more accurate, and sometimes less accurate, than individual decisions (Reagan-Cirincione, 1994). However, the kind of consensus effect referred to in the present context concerns whether decisions are more accurate when there is an agreement amongst independently derived judgements. According to the general principle of aggregation, we might expect aggregated judgements to be more reliable and valid than individual scores. Perhaps the closest examples we have here are from the area of consensus forecasts (of weather, economy, production etc.), which, due to the principle of aggregation, consistently tend to be better when there is a consensus amongst independently derived forecasts (Blix, Wadefjord, Wienecke & Adahl, 2001).

Although previous findings on low stakes situations have found little variance in observer accuracy in deception detection (Bond & DePaulo, 2008; Levine, 2010), this would perhaps not be surprising if there were few discernible cues to detect. However, if there are discernible cues to detect, it could be that consensus judgements derived from independent judges will most enable us to identify them accurately.

### **11.1.1. Review of Aims and Hypotheses**

In light of the above considerations, the main aim of Study 7 was to investigate the comparative accuracy of police and non-police participants, in detecting deception in the real life, high stakes situation of appeals. So far, only one study (Vrij & Mann, 2001a) using police officers as observers has looked at appeals; in this, detection accuracy was not above chance; however, only five appeals were used, all of which were deceptive, and they were not in the native language of the observers. The choice of police officers in the present study into observer accuracy was guided not only by the fact that they have formed the main observer samples in high stakes lies, but by the obvious fact that, in forensic contexts, the consequences of faulty credibility judgements can be particularly serious.

In Studies 2 and 4 of the present thesis, participants from the general public were able to detect deception above chance levels (at 75% and 67%, respectively); however, as the primary aim of those studies was to investigate subjective observer ratings of cues to deception, not observer accuracy, only a small number of participants were used, and it is unclear how observers would respond without first rating the appeals on the cues (i.e. in these studies, the observers first rated the cues, and then made an overall credibility judgement). The present study, therefore, investigated accuracy in detecting deception in a larger group of non-police observers, who did not rate cues to deception in the appeals prior to making credibility judgements. It was hypothesised, nevertheless, that, in general,

participants, both police and non-police, would produce accuracy rates in detecting deception above chance, because of the high stakes nature of the stimulus materials.

However, it was further hypothesised that, due to the likelihood of increased exposure to deceptive behaviour in the course of their jobs (domain familiarity), police officers would be more accurate at detecting deception than the general public. The basic idea here is that continued exposure to individuals who are lying, and have subsequently been ‘proved’ to have been lying (in a legal sense), would enable observers more readily to detect cues that differentiate liars from truth-tellers. By the same token, it was also expected that, within the police sample, officers in CID, who routinely interview suspects, would be more accurate in judging credibility than firearms officers, who do not routinely interview suspects.

Also, as in studies 2 and 4, it was further hypothesised that, because in this high stakes situation cues to deception would be more likely to emerge than in low stakes situations, observers would be less likely to rely on a truthfulness heuristic (i.e. simply assume the person is telling the truth), thus eliminating any truth bias. Indeed, given the greater exposure of police officers to high stakes deception in their jobs, there was even a possibility of a deception bias in the police sample.

Additionally, despite mixed findings in previous research, it was hypothesised that the use of real life, high stakes deception as stimulus materials, would result in a positive relationship between accurate veracity judgements and confidence in all participants, and, due to the principle of aggregation, consensus judgements would accurately predict veracity. A further aim of the present study was

to investigate whether appealers who were accurately identified as honest or deceptive by observer consensus, also exhibited the cues identified in Studies 4, 5 and 6 as reliably discriminating between honesty and deception. As the detection of deception must necessarily rely on the presence of cues to deception, it was hypothesised that appeal scores in the case by case analyses in Studies 4, 5 and 6, would be related to accurate observer judgements of the appeals in the present study.

In the present study it was also expected that there would be large variations in accuracy between individuals within populations. Although some previous studies have shown some variability in observer accuracy (Ekman & O'Sullivan, 1991; Vrij & Graham, 1997; Vrij & Mann, 2001a; Vrij & Mann, 2001b), they have typically used a limited range of stimulus materials (between one and 10 'senders' in the studies mentioned), so that findings regarding individual differences in accuracy may not be robust. The present study, therefore, used a relatively large number of stimulus materials (36) in an attempt to address this issue. Related to this issue, a further aim of the present study was to investigate whether the most accurate observers used the cues identified in Studies 4, 5 and 6 as discriminating between honesty and deception in this context. It was tentatively expected that accurate observers would rely on some of the cues identified in the present thesis more often than low accuracy observers.

## **11.2. Method**

### **11.2.1. *Participants***

A total of 107 participants were recruited to take part in the study. The participants were 37 firearms officers, and 33 CID officers, of a large North West police service, and 37 undergraduate students participating for course credit. Students were recruited using the University online participation system, and police officers were recruited using opportunity sampling via the researcher's contact with a training sergeant at a large North West police service. All participation was voluntary. There were 33 male and four female firearms officers, ages ranged from 26 to 53 ( $M = 37.27$ ,  $SD = 6.98$ ), and years of experience as a police officer ranged from six to 29 years ( $M = 13.35$ ,  $SD = 6.38$ ). There were 18 male and 15 female CID officers, whose ages ranged from 26 to 46 years ( $M = 34.18$ ,  $SD = 6.40$ ), and years of experience as a police officer ranged from two to 23 years ( $M = 9.67$ ,  $SD = 5.73$ ). There were four male and 33 female undergraduate students; their ages ranged from 18 to 25 years ( $M = 19.19$ ,  $SD = 1.58$ ). None of the undergraduate students had been a police officer.

### **11.2.2. *Materials and Procedure***

Participants attended the study in a psychology laboratory at the University, in groups ranging in size from three participants to 14 participants. There were seven



groups of firearms officers, three groups of CID officers, and four groups of students. This variation in group size was a result of the number of police officers able to participate at any particular session, and did not affect the running of the study. Each participant was given an instruction and response sheet. Participants were informed that they would be viewing short video clips of people appealing for help with missing or murdered relatives, and would be asked to decide whether they thought each appellant was lying (i.e. was actually involved in the death or disappearance of their relative), or telling the truth (i.e. was not involved in the death or disappearance of their relative). For each appeal, participants were also asked to rate how confident they were in their credibility judgement, on a Likert scale of 1 (not at all confident) to 5 (very confident). Before each appeal was shown, a short summary was provided of what was publicly known of the case at the time of the appeal, for example, 'This clip features a man talking about his missing 34 year old daughter, Nancy'. Each summary included: the relationship between the appellant and the relative (e.g. parent, spouse, sibling, child); the age of the missing or murdered relative if the appellant was a parent (to allow for possible differences in observers' expectations of parental behaviour depending on whether, for example, their child was an infant or an adult); an explanation of any names, events or details mentioned in the appeal; and whether the relative was missing (i.e. no body had been found) or publicly known to be dead (i.e. the body had been found). Participants were asked to check a box on the response sheet if they were familiar with the appellant featured in the clip, or the outcome of the case, and to not complete the section of the response sheet for

that appeal. Before any appeals were shown, participants were asked to estimate the percentage of the appeals that they thought they would correctly identify as honest or deceptive.

The appeals used in the present study were the 32 appeals used in Studies 4, 5 and 6 of the present thesis, 16 of which were honest, and 16 of which were deceptive, plus four extra appeals. A description of the appeals and criteria for selection is provided in Chapter 5. See Appendix 3 for case summaries of all the appeals. The extra appeals (two honest and two deceptive) were originally included as possible fillers shown before the other appeals; in a small pilot study it was found that accuracy was low for the first few appeals shown, suggesting that it may be beneficial for observers to have some familiarity with the type of stimulus materials used in the study. However, in the present study, as accuracy for the filler appeals was comparable to accuracy for the rest of the appeals, they were included in further analyses, giving a sample size of 36 appeals. In the group of firearms officers, the number of appeals with which any individual was familiar with ranged from zero appeals (four participants) to six appeals (one participant),  $M = 2.43$ ,  $SD = 1.39$ . In the group of CID officers, the number of appeals with which any individual was familiar with ranged from zero appeals (one participant) to six appeals (two participants),  $M = 3.45$ ,  $SD = 1.62$ . In the group of students, the number of appeals with which any individual was familiar with ranged from zero appeals (15 participants) to five appeals (two participants),  $M = 1.05$ ,  $SD = 1.37$ .

For each group, the appeals were shown in a different, randomised order, except for the first four appeals which, as noted above, were always shown first. Participants were shown the first of the 36 appeals, and then completed the relevant section of the response sheet. This process was repeated for each appeal. After viewing all 36 appeals, participants were asked to note what percentage of the appeals they thought that they had correctly judged on veracity.

Participants then completed the second part of the response sheet, on which they were asked to indicate, on a Likert scale of one to five (from 1 never, to 5 very often), how often they used certain cues to come to their veracity decisions. If they used a cue, they were asked to indicate whether they thought it indicated honesty or deception. Thirty five cues were included in this section. Study 4 (investigating subjective cues to deception) was conducted concurrently with the present study, and as results were not available regarding which cues discriminated between honest and deceptive appeals, all 28 cues investigated in Study 4 were used in this section of the present study. Seven cues that had discriminated between honest and deceptive appeals (on a one or two-tailed test) from Study 5 (frequency counts of verbal behaviours) were also included; equivocal language, speech errors, avoiding the use of brutal language, expressions of concern, grief or pain, expressions of positive emotion towards the relative, references to norms of emotion, and hope of finding a missing relative alive. However, two of the three cues investigated in Study 6 (frequency counts of nonverbal cues) were not included in the present study, as again, because of constraints on the availability of the police participants, Study 7

was conducted before cues to be included in Study 6 had been finalised. See Appendix 11 for an example of the full response sheet. Participation in the study lasted between one and a half, and two hours.

### **11.3. Results**

#### **11.3.1. Accuracy in Detecting Deception, and Truth Bias**

To investigate the hypothesis that participants would produce accuracy rates in detecting deception above chance, overall accuracy rates were calculated for each participant. Any cases which were known to a participant were not included in the accuracy scores (the majority of participants were not familiar with more than two of the appealers, and none was familiar with more than six). Overall mean accuracy was 70.75% ( $SD = 8.68$ ), median accuracy was 71.88 %.

As hypothesised, all groups of participants produced percentage accuracy rates above chance. Firearms officers achieved a median accuracy of 72.79%, which a sign test showed to be significantly above chance,  $Z = 5.92$ ,  $p < .001$  (all values were above 50). CID officers achieved a median accuracy of 74.17%, which a sign test showed to be significantly above chance,  $Z = 5.57$ ,  $p < .001$  (all values were above 50). The public achieved a median accuracy of 67.65%, which a sign test

showed to be significantly above chance,  $Z = 5.50$ ,  $p < .001$  (35 values were above 50).

To investigate whether there were differences in accuracy between the three groups of participants, and also between veracity conditions (i.e. deceptive appeals and honest appeals), a 3 x 2 mixed ANOVA (group: firearms, CID and public x veracity: deceptive and honest), with repeated measures on the second factor was conducted on the percentage accuracy scores; the means and *SDs* are shown in Table 11.1.

A significant main effect was found for group,  $F(2,104) = 3.37$ ,  $p = .038$ ,  $\eta^2_p = .061$ . Post hoc tests revealed that, as hypothesised, firearms officers ( $p = .031$ ) and CID officers ( $p = .024$ ) were significantly more accurate than the public. However, the difference in accuracy between firearms officers and CID officers was not significant ( $p = .874$ ), offering no support for the hypothesis that CID officers would be more accurate at detecting deception than firearms officers. See Table 11.1.

It can be noted that not only were police officers more accurate than the public sample used here, they were also older (mean age of police officers was 35.81 years, whereas the mean age of the student public was 19.19 years). One way of accounting for this would obviously have been to have included age as a covariate in the accuracy analyses and performed ANCOVAs. However, because of the nature of the age difference between the samples (there was no overlap between the ranges), this would have spuriously reduced any differences between the samples not due to the age variable. Instead, therefore, a different approach was adopted to assess the

potential confounding effect of age; Pearson's correlations were performed to assess whether accuracy was related to age within the samples. For police officers, accuracy was not related to age,  $r = -.05$ ,  $p = .708$ , or to years of experience as a police officer,  $r = -.09$ ,  $p = .461$ . And, in the student sample, again accuracy was not related to age,  $r = 0.24$ ,  $p = .155$ . Indeed, as can be seen, within each sample the correlations between age and accuracy were low and not significant. This means it is extremely unlikely that the results for accuracy were confounded by age per se.

There was no significant main effect for veracity,  $F < 1$ ,  $p = .474$ ,  $\eta^2_p = .005$ , suggesting initial support for the hypothesis that truth bias would be eliminated. However, there was a significant interaction between group and veracity,  $F(2,104) = 9.97$ ,  $p < .001$ ,  $\eta^2_p = .161$ . Post hoc tests showed that firearms officers' accuracy did not differ significantly between deceptive and honest appeals, although there was a trend ( $p = .089$ ) for firearms officers to be more accurate when judging deceptive appeals than when judging honest appeals. CID officers were significantly more accurate ( $p = .007$ ) when judging deceptive appeals than when judging honest appeals. These findings were in line with the expectation that police officers may show some evidence of deception bias. The public were significantly more accurate ( $p = .002$ ) when judging honest appeals than when judging deceptive appeals, suggesting that, in the case of the public, the hypothesis that truth bias would be eliminated was not supported.

Further post hoc  $F$  tests revealed that firearms officers ( $p < .001$ ), and CID officers ( $p < .001$ ), were more accurate than the public when judging deceptive

appeals. There was no difference in accuracy between firearms officers and CID officers when judging deceptive appeals ( $p = .922$ ). When judging honest appeals, there were no significant differences in accuracy between firearms officers and the public ( $p = .258$ ), CID officers and the public ( $p = .191$ ), or firearms officers and CID officers ( $p = .975$ ). See Table 11.1.

Table 11.1: *Accuracy of observers' credibility judgements of appeals*

	<u>Deceptive Appeals</u>	<u>Honest appeals</u>	<u>All appeals</u>
	<i>M</i> (SD)	<i>M</i> (SD)	<i>M</i> (SD)
<b>Firearms Officers</b>	75.55% (13.07)	68.73% (15.98)	72.14% (8.49)
<b>CID officers</b>	76.86% (13.10)	68.07% (10.42)	72.47% (7.98)
<b>Public</b>	62.21% (16.09)	73.43% (10.89)	67.82% (8.92)
<b>All participants</b>	71.34% (15.57)	70.15% (12.87)	70.75% (8.68)

### 11.3.2. Accuracy and Confidence in Credibility Judgements

Participants' general confidence was measured before the study (they were asked to note what percentage of the appeals they thought they would judge correctly), and

after the study (they were asked to note what percentage of the appeals they thought they had judged correctly). A 3 x 2 mixed ANOVA (group: firearms, CID and public x time: pre-test and post-test), with repeated measures on the second factor was conducted on the percentage confidence scores; the means and *SDs* for the significant effects are shown in Table 11.2.

Table 11.2: *Observers' self-estimates of pre & post test accuracy, and confidence rated on a 5 point scale in their credibility judgements of appeals*

	<b>Firearms officers</b>	<b>CID officers</b>	<b>Public</b>	<b>All participants</b>
Pre-test participant self-estimate of accuracy <i>M</i> ( <i>SD</i> )	64.76% (9.48)	64.58% (11.75)	60.86% (12.70)	63.36% (11.41)
Post-test participant self-estimate of accuracy <i>M</i> ( <i>SD</i> )	54.03% (12.88)	49.55% (15.23)	44.78% (15.38)	49.45% (14.89)
<b><u>Deceptive appeals</u></b>				
Correct judgement confidence rating <i>M</i> ( <i>SD</i> )	3.73 (0.53)	3.42 (0.58)	3.07 (0.65)	3.41 (0.65)
Incorrect judgement confidence rating <i>M</i> ( <i>SD</i> )	3.15 (0.63)	2.64 (0.74)	2.65 (0.58)	2.82 (0.69)
Overall confidence rating <i>M</i> ( <i>SD</i> )	3.62 (0.50)	3.22 (0.58)	2.91 (0.55)	3.25 (0.61)
<b><u>Honest appeals</u></b>				
Correct judgement confidence rating <i>M</i> ( <i>SD</i> )	3.43 (0.48)	2.89 (0.59)	3.13 (0.63)	3.16 (0.61)
Incorrect judgement confidence rating <i>M</i> ( <i>SD</i> )	3.16 (0.74)	2.90 (0.77)	2.58 (0.71)	2.87 (0.77)
Overall confidence <i>M</i> ( <i>SD</i> )	3.37 (0.47)	2.87 (0.58)	3.01 (0.58)	3.09 (0.58)



<b><u>All appeals</u></b>				
Correct judgement confidence rating <i>M</i> (SD)	3.62 (0.45)	3.16 (0.52)	3.12 (0.57)	3.30 (0.56)
Incorrect judgement confidence rating <i>M</i> (SD)	3.17 (0.54)	2.74 (0.69)	2.65 (0.57)	2.86 (0.64)
Overall confidence <i>M</i> (SD)	3.50 (0.44)	3.04 (0.54)	2.96 (0.53)	3.17 (0.55)

---

A significant main effect was found for time,  $F(1,104) = 106.84$ ,  $p < .001$ ,  $\eta^2_p = .507$ , indicating that participants were more confident of their overall accuracy before participating in the study, than after participating in the study. A significant main effect was also found for group,  $F(1,104) = 3.37$ ,  $p = .038$ ,  $\eta^2_p = .061$ . Post hoc tests revealed that firearms officers were more confident than the public ( $p = .012$ ). There was no difference in confidence between firearms officers and CID officers ( $p = .379$ ), or between CID officers and the public ( $p = .111$ ). The interaction between time and group was not significant,  $F(1,104) = 1.52$ ,  $p = .224$ ,  $\eta^2_p = .028$ .

To investigate whether accuracy in credibility judgements was related to confidence in credibility judgements, mean confidence ratings were calculated for each participant for correct judgements of deceptive appeals, correct judgements of honest appeals, incorrect judgements of deceptive appeals, and incorrect judgements of honest appeals. A 3 x 2 x 2 mixed ANOVA with repeated measures on the last two factors was conducted on the mean confidence ratings, with the factors group (firearms, CID and public), veracity (deceptive appeal and honest appeal), and

judgement accuracy (correct judgement and incorrect judgement). The means and *SDs* are presented in Table 11.2.

As hypothesised, a significant main effect was found for judgement accuracy,  $F(1,102) = 126.11, p < .001, \eta_p^2 = .553$ , such that, overall, participants were more confident in their correct judgements, than in their incorrect judgements. A significant main effect was also found for veracity,  $F(1,102) = 6.00, p = .016, \eta_p^2 = .056$ , so that overall, participants were more confident in their judgements of deceptive appeals, than in their judgements of honest appeals.

A significant main effect was also found for group,  $F(2,102) = 9.41, p < .001, \eta_p^2 = .156$ . Post hoc tests showed that overall, firearms officers were more confident in their judgements than CID officers ( $p = .002$ ), and the public ( $p < .001$ ), lending partial support to the hypothesis that police participants would be more confident in their judgements than the public.

A significant interaction was found between veracity and judgement accuracy,  $F(1,102) = 13.28, p < .001, \eta_p^2 = .115$ . Post hoc tests showed that for both honest and dishonest appeals, participants were more confident when making a correct judgement, than when making an incorrect judgement ( $p < .001$ ). However, the interaction was caused by a difference in confidence between correct judgements for honest and dishonest appeals. Confidence was significantly higher for correct judgements in deceptive appeals, than for correct judgements in honest appeals ( $p < .001$ ), whereas this was not the case for incorrect judgements.

A significant interaction was also found between veracity, judgement accuracy, and group,  $F(1,102) = 9.34$ ,  $p < .001$ ,  $\eta^2_p = .155$ . Post hoc tests showed that the interaction was caused by the CID officers' confidence ratings for honest appeals. For all groups, for all appeals, confidence ratings were significantly higher for correct judgements than for incorrect judgements ( $p < .001$ ), except for CID confidence ratings on honest appeals, for which there was no significant difference in confidence ratings between correct and incorrect judgements. Again see Table 11.2.

### ***11.3.3. Relationship of Consensus Judgements to Veracity***

To reiterate, the idea behind the consensus hypothesis is that when there is high consensus as to the veracity of a particular appeal, it is more likely to be accurate than inaccurate. So, for example, if 75% of people agree that a particular appeal is dishonest, it is most likely dishonest. To investigate the utility of consensus judgements in determining the veracity of the appeals, therefore, for each appeal, the percentage of correct judgements of veracity ('hits'), was compared with the percentage of incorrect judgements of veracity ('misses'), using Wilcoxon signed ranks tests. Hence this analysis looks at agreement on appeals, not whether individual observers are accurate.

As expected, the percentage of correct judgements by participants (median = 72.77), was significantly greater than the percentage of incorrect judgements (median

= 27.23),  $Z = -4.76$ ,  $p = .001$ . Indeed, in 33 out of the 36 appeals (92%), the percentage of ‘hits’ was greater than the percentage of ‘misses’. A consensus of 75% or over was always correct, and 17 appeals had a consensus over 75%. See Table 11.3. In other words, the more likely, collectively, people were to agree on the direction of a judgement, the more correct they were.

Looking at police and public consensus separately, the percentage of ‘hits’ by police participants (median = 81.00) was significantly greater than the percentage of ‘misses’ by police participants (median = 19.00),  $Z = -4.80$ ,  $p < .001$ , and in 33 out of 36 appeals, the percentage of police participant ‘hits’ was greater than the percentage of police participant ‘misses’. The percentage of ‘hits’ by public participants (median = 68.00) was significantly greater than the percentage of ‘misses’ by public participants (median = 32.00),  $Z = -4.28$ ,  $p < .001$ , and in 29 out of 36 appeals, the percentage of public participant ‘hits’ was greater than the percentage of public participant ‘misses’. See Table 11.3.

In deceptive appeals, the percentage of ‘hits’ (median = 74.38) was significantly greater than the percentage of ‘misses’ (median = 25.62),  $Z = -3.33$ ,  $p = .001$ , and in 16 out of 18 deceptive appeals, the percentage ‘hits’ was greater than the percentage of ‘misses’. In deceptive appeals, the percentage of ‘hits’ by police participants (median = 84.00) was significantly greater than the percentage of ‘misses’ by police participants (median = 16.00),  $Z = -3.45$ ,  $p = .001$ , and in 16 out of 18 deceptive appeals, the percentage of police participant ‘hits’ was greater than the percentage of police participant ‘misses’. In deceptive appeals, the percentage of

‘hits’ by public participants (median = 65.50) was significantly greater than the percentage of ‘misses’ by public participants (median = 34.50),  $Z = -2.31$ ,  $p = .021$ , and in 12 out of 18 deceptive appeals, the percentage of public participant ‘hits’ was greater than the percentage of public participant ‘misses’. See Table 11.3.

In honest appeals, the percentage of ‘hits’ (median = 68.77) was significantly greater than the percentage of ‘misses’ (median = 31.23),  $Z = -3.46$ ,  $p = .001$ , and in 17 out of 18 honest appeals, the percentage of ‘hits’ was greater than the percentage of ‘misses’. In honest appeals, the percentage of ‘hits’ by police participants (median = 74.50) was significantly greater than the percentage of ‘misses’ by police participants (median = 25.50),  $Z = -3.38$ ,  $p = .001$ , and in 17 out of 18 honest appeals, the percentage of police participant ‘hits’ was greater than the percentage of police participant ‘misses’. In honest appeals, the percentage of ‘hits’ by public participants (median = 70.00) was significantly greater than the percentage of ‘misses’ by public participants (median = 30.00),  $Z = -3.53$ ,  $p < .001$ , and in 17 out of 18 honest appeals, the percentage of public participant ‘hits’ was greater than the percentage of public participant ‘misses’. See Table 11.3.

To investigate whether accurate consensus judgements were related to the length of the appeal, (i.e. it may be that a high level of accurate consensus is related to longer exposure to the ‘sender’), a Pearson’s correlation was conducted on the length of the appeals (seconds), and the number of ‘hits’ on the appeal. The relationship between length of appeal and percentage of correct judgements made about the appeal was not significant,  $r = .11$ ,  $p = .513$ .

Table 11.3: *For each appeal; percentage of accurate judgements (overall, by police, and by the public), and percentage of 'hits' on cues identified in Studies 4, 5 and 6 as discriminating between honest and deceptive appeals.*

	Overall 'hits'	Police 'hits'	Control 'hits'	Case 'hits' on cues rated subjectively by observers	Case 'hit' on cues measured by frequency counts
<b><u>Deceptive appellers</u></b>					
Case 10	87%	84%	92%	94%	80%
Case 11	80%	84%	73%	100%	80%
Case 12	96%	99%	92%	100%	80%
Case 13	29%	24%	38%	35%	60%
Case 14	85%	86%	83%	71%	80%
Case 15	71%	71%	74%	94%	50%
Case 16	43%	48%	35%	12%	60%
Case 17	82%	90%	78%	100%	83%
Case 18	73%	84%	51%	94%	40%
Case 19	63%	76%	38%	29%	100%
Case 20	54%	54%	54%	0%	83%
Case 21	62%	67%	49%	0%	67%
Case 22	63%	78%	34%	0%	50%
Case 23	82%	86%	76%	100%	17%
Case 24	83%	91%	68%	100%	83%
Case 25	76%	92%	63%	100%	100%
Case 6	92%	96%	84%		
Case 9	73%	86%	47%		

	Overall 'hits'	Police 'hits'	Control 'hits'	Case 'hits' on cues rated subjectively by observers	Case 'hit' on cues measured by frequency counts
<b><u>Honest appealers</u></b>					
Case 26	53%	52%	54%	53%	60%
Case 27	56%	51%	65%	65%	50%
Case 28	81%	81%	81%	88%	67%
Case 29	89%	93%	86%	100%	100%
Case 30	86%	83%	92%	88%	75%
Case 31	59%	53%	68%	24%	20%
Case 32	64%	53%	87%	100%	100%
Case 33	66%	69%	62%	100%	75%
Case 34	64%	61%	68%	41%	20%
Case 35	81%	80%	84%	47%	60%
Case 36	82%	87%	72%	65%	100%
Case 37	90%	86%	97%	100%	60%
Case 38	62%	59%	68%	29%	67%
Case 39	71%	81%	63%	82%	83%
Case 40	93%	90%	97%	100%	100%
Case 41	38%	38%	37%	94%	83%
Case 42	83%	83%	84%		
Case 5	52%	52%	53%		

#### **11.3.4. Relationship Between Correct Veracity Judgements and ‘hits’ on Cues**

To investigate the hypothesis that the percentage of ‘hits’ (i.e. correct veracity judgements) on each appeal, would be related to the number of ‘hits’ on the cues identified in Studies 4, 5 and 6, as discriminating between honest and deceptive appeals, a further series of Pearson’s correlations was conducted. The four filler appeals (cases 5, 6, 9 and 42) were not included in these analyses as they were not coded in Studies 4, 5 and 6.

As hypothesised, the correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ on the subjective cues (in Study 4), was significant,  $r = .59$ ,  $p < .001$ . However, the correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ measured by frequency counts (in Studies 5 and 6), was not significant,  $r = .26$ ,  $p = .156$ . See Table 11.3.

For deceptive appeals, the correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ on the subjective cues (in Study 4), was significant,  $r = .74$ ,  $p = .001$ . However, the correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ measured by frequency counts (in Studies 5 and 6), was not significant,  $r = .14$ ,  $p = .595$ . See Table 11.3.

For honest appeals, the correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ on the subjective cues (in



Study 4), was not significant,  $r = .36$ ,  $p = .170$ , and the correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ measured by frequency counts (in Studies 5 and 6), was also not significant,  $r = .37$ ,  $p = .158$ . See Table 11.3.

#### ***11.3.5. Relationship Between Observer Accuracy and Observer Cue Use***

As expected, there were large variations in individual accuracy, which ranged from 91% to 45%. To investigate whether the most accurate observers were most likely to use the cues identified in Studies 4, 5 and 6 as discriminating between honesty and deception in this context, participants were grouped into quartiles according to their overall percentage levels of accuracy; low (45% - 64%), medium low (65% - 71%), medium high (72% - 76%), and high (77% - 91%). It can be noted that the high accuracy group included 20 police officers and six participants from the public, and also that 50% of police officers achieved accuracy rates of 74% or higher (correspondingly, only 27% of public participants achieved accuracy rates of 74% or higher).

A part of the procedure, participants had indicated how often they used each of the 35 cues included on the response sheet, and if they used it, whether they thought that it indicated honesty or deception. This was coded on a scale of 4 to -4, such that cues used very often in the expected direction had a score of 4, cues that

were never used had a score of 1, and cues that were used often but in the opposite direction to the one expected had a score of -4.

It will be remembered that this study was conducted concurrently with Studies 4 and 6, such that, at the time, it was not known which of the 35 cues would actually discriminate successfully between honest and deceptive appeals. Hence the following analysis was conducted only on the 22 cues subsequently shown to discriminate in the other studies (positive emotion expressed towards the relative, references to norms of emotion/behaviour, expressions of hope, equivocation, gaze aversion, fake emotion, fake facial expression, putting on a performance, dislike for the appealer, creepy, behaving unnaturally, no sympathy for the appealer, genuinely sad, sad eyes, genuine and heartfelt appeal, feel the appealer's pain, feel sorry for the appealer, normal, urgency, personal and expressive voice, plausible, and voice quivering with emotion). To investigate whether there were differences between the groups (low accuracy, medium low accuracy, medium high accuracy, and high accuracy), in how they reported using the 22 discriminatory cues, a MANOVA with percentile accuracy group as the between subjects independent variable, was conducted on the scores for reported cue use. Using Pillai's Trace the multivariate test was not significant,  $F(15,303) = 1.51, p = .101$ . In other words, there was no evidence that the accurate observers were more likely to report using the cues identified in the previous studies in this thesis when making their judgments of veracity.

#### **11.4. Discussion**

The general hypothesis that all participants (both police and public) would achieve accuracy rates in detecting deception above chance, due to the high stakes nature of the stimulus materials, was supported. Overall mean accuracy was 71%, and all groups (firearms, CID, and public) achieved mean accuracy scores above chance levels (at 72%, 72%, and 68% respectively), much higher than the meta-analytic average of 54% reported from mainly low stake studies (Bond & DePaulo, 2006). This is a particularly important finding, as although it has been reported in previous research (Mann & Vrij, 2006; Mann et al., 2004, 2006; Vrij & Mann, 2001b; Vrij et al., 2006) that police officers are able to achieve accuracy in line with the present findings, when judging high stakes deception (at 68%, 65%, 69%, 64% and 72% respectively), the ability of non-police participants to do the same has not been previously investigated. Moreover, as the detection of deception necessarily implies differences between honest and deceptive behaviour, these findings support theoretical proposals that increasing the stakes may exacerbate factors underlying the production of cues to deception, resulting in more, and/or more prominent, cues to deception, which are discernible to observers.

The hypothesis that police officers would be more accurate at detecting deception than the public, was also supported. This suggests that findings from previous research, which have reported that police participants are no more accurate than non-police participants (for example, DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991), may have occurred at least partly because of the low stakes

contexts of the stimulus materials. In the present study, not only were the police participants more accurate overall than the public, but half of all police participants achieved accuracy rates of 74% or more, whereas only 27% of non-police participants achieved this level of accuracy. It was suggested that one possible reason why police observers might be more accurate than the public in detecting deception, is that they would have been more exposed to deceptive behaviour in the course of their jobs; however, the finding that there was no relationship between years of experience as a police officer, and accuracy, casts some doubt on this as an explanation. It may be that police officers have a greater expectation of deception; this will be explored more fully in the general discussion.

Considering that meta-analytic reports (mainly involving low stakes studies) indicate that people are no more confident when making accurate veracity decisions than when making inaccurate veracity decisions, the finding that there was a positive relationship between accurate veracity judgements and confidence, is potentially an important one. This finding endorses the hypothesis that the use of high stakes stimulus materials, in which cues to deception may be more prominent, has an effect not only on observer accuracy in detecting deception, but also on the appropriate calibration of confidence. Moreover, although participants were more confident in their correct decisions than their incorrect decisions for both honest and deceptive appeals, they were more confident in their correct decisions for deceptive appeals, than in their correct decisions for honest appeals. This is an interesting finding; if observers are more confident when correctly identifying deceptive appeals

than honest appeals, it implies that deceptive appeals may be more ‘obvious’ to observers, than honest appeals. This fits with other more general research on witness confidence and accuracy indicating that confidence-accuracy relationships tend to be higher when the items to be calibrated include a substantial number that are ‘obvious’, because such items increase the variance (Kebbell, Wagstaff & Covey 1996; Wheatcroft, Wagstaff, & Kebbell, 2004). The relevance of this to other findings in the present study, and to other studies in the present thesis, is also discussed shortly.

Based on previous research, it was hypothesised that police participants would be more confident in their judgements overall, than participants from the public, and this hypothesis was only partially supported. Firearms officers were found to be more confident in their decisions than either the public or CID officers, whereas CID officers were no more confident than the public. Speculatively, the overall higher confidence of the firearms officers may be due to aspects of the selection procedure for their specialism, which may favour more confident personalities. Interestingly, although overall participant accuracy was high, and participants were more confident in their correct judgements than in their incorrect judgements, their overall confidence in their ability to detect deception was low; all groups of participants’ estimates of their own accuracy after completing the study hugely underestimated their actual performance (firearms officers’ mean estimate of their own accuracy was 54%, when they actually achieved 72% mean accuracy; CID officers’ mean estimate of their own accuracy was 50%, when they actually achieved

72% mean accuracy; and the public's mean estimate of their own accuracy was 45%, when they actually achieved 68% mean accuracy). Moreover, participants' estimates of their own accuracy was lower after completing the study, than in anticipation of the study, suggesting that they did not find the task easy. In contrast to previous indications that police officers are poor at detecting deception (for example, DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991; Meissner & Kassin, 2002; Vrij & Graham, 1997), and yet are over-confident in their ability to do so (for example, DePaulo & Pfeifer, 1986; Meissner & Kassin, 2002), the present findings suggest that police officers (and the public), are able to detect high stakes deception well above the levels of chance, but underestimate their own ability to do so.

The investigation of consensus judgements in the present study was particularly important, as this is an aspect of deception detection that has not been previously explored. It was expected that, due to the high stakes nature of the stimulus materials, and the likelihood that this would result in more, and more discernible cues to deception, an accurate consensus would emerge for some appeals. However, the finding that 33 out of the 36 appeals (92%) could be correctly classified as honest or deceptive using consensus judgments, was not expected. It is extremely unusual in deception research for any variable to predict veracity with accuracy as high as this; in the present thesis, the cues to deception most predictive of veracity were gaze aversion (79%), creepy, no sympathy for the appellant, and voice quivering with emotion (all 75%), none of which was as predictive as consensus. Evidently, as the utility of consensus judgements has not been previously

investigated, this finding can only be regarded as preliminary, and replication would be needed before any firm conclusions can be drawn. For example, future research investigating the value of consensus judgements in a different high stakes context would be useful. Nonetheless, the present findings suggest that, in the context of appeals at least, there are sufficient deceptive and honest behaviours produced by the ‘senders’, for the majority of observers to make correct veracity judgements about them. It appears that Levine’s idea that above chance detection of deception is due to a few ‘leaky liars’ (2010) may not apply in this high stakes context, as nearly all appealers (both deceptive and honest) were sufficiently ‘leaky’ to be correctly classified by the consensus of observers.

The hypothesis that appealers that were accurately identified as honest or deceptive by observer consensus, would also exhibit the cues identified in Studies 4, 5 and 6 as reliably discriminating between honesty and deception, was only partially supported. There was no significant relationship between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ on the verbal and nonverbal cues measured by frequency counts in Studies 5 and 6. This suggests that the participants may not have been using these cues in making their veracity judgements, and thus there may be useful information that is available in the appeals (those verbal and nonverbal cues measured by frequency counts), which observers were not generally using.

However, specifically for the deceptive appeals, there was a strong correlation between the percentage of correct veracity judgements on each appeal,

and the percentage of ‘hits’ on the subjective cues rated by observers in Study 4; appeals that were classified as deceptive using subjective measures, were also correctly identified as deceptive by observer consensus in the present study, although this was not the case for honest appeals. As initially suggested in the analysis of Study 1 of the present thesis, it may be that deceptive behaviour violates certain norms which go unnoticed when conformed to, but are noticed when violated, and result in behaviour that is regarded as aberrant, and that elicits negative personal reactions. In support of this proposition, the findings of Study 4 of the present thesis suggest that liars may exhibit characteristics which appear aberrant (for example, creepy, unnatural, and not normal), and which elicit negative personal reactions (for example, dislike, no sympathy, and a lack of feeling sorry for the appealer, or feeling the appealer’s pain). The suggestion that norms of behaviour may be noticed when they are violated, but when they are conformed to go unnoticed, is in line with the finding that the percentage of ‘hits’ on the subjective cues was related to correct veracity judgements of deceptive appeals, but not of honest appeals.

Furthermore, as discussed above, participants were more confident in their correct judgements of deceptive appeals, than in their correct judgements of honest appeals, suggesting that deceptive appeals may be more ‘obvious’ to observers than honest appeals. Moreover, despite large differences in the length of appeals (some were less than 20 seconds, some were over two minutes long), there was no relationship between the percentage of correct credibility judgements, and the length of the appeal. This implies that, as providing participants with longer examples of



behaviour did not affect their accuracy, participants may have been making rapid, accurate, global judgements, rather than relying on careful analysis of lengthy sections of behaviour. This suggests that participants may have been looking at a ‘type’ of person, rather than specific cues; this is in line with the finding that, for the deceptive appeals at least, correct veracity judgements were related to ‘hits’ on the subjective cues, but not on the verbal and nonverbal cues measured by frequency counts; global impressions may be more readily captured by subjective measures, than by frequency counts of behaviours. The theoretical implications of this are more fully addressed in the general discussion.

As expected, there were large variations in accuracy between individuals, and some individuals were highly accurate. However, the most accurate observers did not reportedly use the cues identified in Studies 4, 5 and 6 as discriminating between honesty and deception in this context, more often than the least accurate observers. This finding raises the question that if the most accurate participants were not using the cues identified in the present thesis, on what cues were they basing their credibility judgements? As previously noted, there have been suggestions in the deception literature that people who are able to accurately detect deception, may use information from across channels of behaviour, and use combinations of behaviour, and behaviours in relation to each other, rather than behaviours in isolation (O’Sullivan & Ekman, 2004); this may explain why in the present study, the most accurate participants did not report relying on individual behaviours more than the least accurate participants. Informal, post-study interviews with three of the most

accurate participants, suggested that they looked for disparity between channels of behaviours, and deviations from norms and expected behaviour, but also that they applied a subtle, nuanced, and often idiosyncratic approach that may be difficult to capture when investigating the use of individual cues.

In sum, the findings from the present study have several important implications. One of the basic propositions of the present thesis was that investigating real life, high stakes deception, in which cues to deception may be more prominent, would produce findings different to investigations of low stake lies in laboratory studies; the findings presented here support this proposition. Contrary to previous findings using low stakes contexts, the present findings suggest that observers (both police and non-police), are able to detect deception well above the levels of chance, and that they are also more confident in their correct credibility judgements than in their incorrect credibility judgements. Furthermore, the findings regarding high consensus in observer credibility judgements are particularly salient, as this aspect of deception detection has not been previously investigated, and yet consensus is unusually highly predictive of veracity.

Indeed, these consensus findings, together with those for individual accuracy, may indicate something highly significant about the use of cues to detect deception in high stakes contexts; that is, simple overall consensus subjective judgements of honesty and dishonesty may equal or even outperform even multiple-cue predictors of deception in predicting who is telling the truth and who is lying. In many respects

this may be one of, if not the most, important of findings in the thesis, and will be discussed further in the final chapter.

## **Part III**

### **General Discussion and Conclusions**

# Chapter 12

## General Discussion and Conclusions

To reiterate, the central focus of the present thesis was to investigate cues to deception and observer accuracy in detecting deception in a real life, high stakes, context-specific situation. The main rationale for this was that, in contrast with a range of largely inconsistent or negative previous findings using a variety of low stakes contexts (Bond & DePaulo, 2006; DePaulo et al., 2003), the use of high stakes, ecologically valid stimulus materials would produce more numerous and more prominent cues to deception, and consequently, observers would be more able to accurately detect deception. Within this general framework, there were four primary aims, which can be reprised as follows.

1. To develop and apply a methodology to identify cues to deception in a specific ecologically valid high stakes context, that of appeals for help with missing and murdered relatives, and to investigate whether these cues are useful in discriminating between honest and deceptive appeals in this context.

2. To develop a multiple-cue approach to behaviours which discriminate between honesty and deception in the context of appeals, and which may be useful in an applied context.
3. To develop a theoretical rationale for cues that may be context-specific and may emerge only in high stakes contexts, and cannot readily be accommodated by existing theoretical approaches.
4. To investigate the effect of using ecologically valid stimulus materials on the accuracy of observers (both police and non-police) in detecting deception, on truth bias, on accuracy-confidence relationships, and the utility of consensus judgements as a predictor of veracity

The aim of this final chapter is to summarize and provide a more general discussion of the outcomes of the present thesis in the context of these aims.

### **12.1. Cues to Deception: Methodological Considerations**

Part of the first main aim of this thesis was to attempt to develop and apply a methodology to identify cues to deception in a specific ecologically valid high stakes context, that of appeals for help with missing and murdered relatives.

### **12.1.1 *Contextual Focus***

As emphasized throughout this thesis, the main methodological rationale for using a specific context was that a range of situational factors may affect the production and salience of cues related to deception, such that all high stakes lies may not necessarily resemble each other in terms of the cues associated with them. However, the small body of literature that has investigated high stakes deception so far has generally lacked a specific contextual focus. This lack of contextual focus may not only have ‘muddied the waters’ regarding which cues discriminate in which contexts, it may also have prevented previously uninvestigated, context-specific cues, from emerging. The findings of the present thesis, and also of two other studies that have used a tightly-focused context (Harpster et al., 2009; ten Brinke & Porter, 2012), suggest that the utilisation of this kind of context specific approach may result in the uncovering of previously unidentified cues. As DePaulo et al. (2003) have pointed out, if such cues exist, it is worth discovering them, even if they are only useful in one context; particularly, when, as here, the context is grave, and the consequences of faulty credibility judgements are very serious.

### **12.1.2. *Between-subjects Design***

Another important methodological consideration in the present thesis was the use of a between-subjects stimulus materials design. As mentioned in the introductory chapters, to date, the majority of high stakes deception studies have utilised within-subjects stimulus materials, where fragments of truthful and deceptive statements from the same individual are compared (for example, Davis et al., 2005; Koper & Sahlman, 1991; Mann & Vrij, 2006; Mann et al.; Mann et al., 2006; Villar et al., 2012; Vrij & Mann, 2001a). However, the application of a between-subjects stimulus design, which compares deceptive with honest individuals, as in the present thesis, may potentially be beneficial for two main reasons. First, it allows context generated cues to deception to emerge that can be identified throughout the deceiver's presentation, even in sections in which he/she is telling the truth. Second, a between-subjects design, allows for cues to emerge which may be related to individual differences, rather than acts of deception per se. In the case of appeals, and in many other high stakes deception contexts, honest and deceptive people differed not only in terms of whether they were lying, but also in terms of whether they had committed serious crimes. There may be differences between the type of person who, for example, murders a relative and then chooses to lie about it, and a person who has not done this; this possible variable is not captured when using a within-subjects design. Moreover, traditional research methods, using experimental, laboratory-based studies, in which participants are randomly allocated to groups, do not allow for possible individual differences to emerge in the type of person who may choose to



deceive, particularly in high stakes situations. As will be discussed shortly, some of the findings of the present thesis are perhaps best accommodated by an individual difference approach of this kind, and could be considered to endorse the utility of this kind of methodology.

### ***12.1.3. Using Cues Generated by Accurate Detectors***

However, perhaps the main methodological innovation in the present thesis was the initial idea of generating cues to deception by examining the reports of individuals previously identified as accurate at detecting deception.

Of the cues originally suggested by the accurate deception detectors in Study 1, 18 discriminated significantly between honest and deceptive appeals in the final analyses (fake emotion, fake facial expression, putting on a performance, equivocation, gaze aversion, dislike for the appealer, creepy, unnatural, lack of sympathy for the appealer, genuine sadness, sad eyes, urgency, genuine/heartfelt, positive emotion expressed towards the relative, hopeful of finding missing relative, normal, feel the appealer's pain, and feel sorry for the appealer). Two of these cues (equivocation and gaze aversion) have been extensively investigated in previous studies, but, as hypothesised, the majority were original and likely to be context-specific (notwithstanding that some of them may be predicted from previous research).

Given these cues were generated by only two individuals, arguably, the present results suggest that investigating the cues used by individuals who are able to accurately detect deception may certainly be a methodology that warrants further development. Furthermore, this kind of inductive approach also resulted in the emergence of an additional previously unidentified cue that was noticed by the researcher (references to norms of emotion or behaviour). In sum, such results suggest that, whilst theory derived ideas are obviously important, a fine-grained, data-driven approach may be particularly useful in this area.

## **12.2. Cues to Deception: Findings**

The second part of this first aim of the thesis was to investigate whether cues generated using the methodologies described in the last section are useful in discriminating between honest and deceptive appeals in this context. To address this further, it is obviously useful to look in more detail at results pertaining to all of the cues investigated in the present work.

Including those identified in Study 1, 23 cues discriminated between honest and deceptive appeals in the final analyses, all with substantial effect sizes, from  $d = 0.73$  (head shaking) to  $d = 1.62$  (expressions of hope). A further three cues discriminated on a one-tailed test (avoid brutal language or detail, expressions of

concern or pain, and speech errors). As noted previously, in DePaulo et al.'s (2003) meta-analysis of cues to deception in low stakes contexts, the median effect size was  $d = 0.10$ , and the most diagnostic cue had a moderate effect size of  $d = 0.66$ . The number of cues identified in the present thesis, and their magnitude, appears to support the basic premise that real life, high stakes deceptive behaviour differs from the type of low stakes deceptive behaviour that is generally examined in experimental, laboratory-based studies.

#### **12.2.1. Findings in Line with Previous Research**

As expected, some cues to deception which discriminated between honest and deceptive appeals in the studies presented here, have also been found to discriminate between honesty and deception in different high stakes contexts; this suggests that some factors underlying the production of cues to deception may be stable across high stakes situations. Moreover, the explanations for these particular findings could be viewed as broadly consistent with the Four Factor theoretical model of deception related behaviours (Zuckerman et al., 1981).

Hence deceptive appealers in the present studies produced more equivocation than honest appealers, and equivocal language has also been related to deception in written statements from suspects/witnesses in police investigations (Adams & Jarvis, 2006), and in another study of appeals (ten Brinke & Porter, 2012). As previously

noted, researchers have suggested that deception may be associated with behaviours that inhibit direct communication, and equivocal language could thus be construed as a form of psychological evasion or distancing from feelings of guilt or shame ((DePaulo et al., 2003; Zuckerman et al., 1981). This may be particularly salient in the context of deceptive appealers, who may use equivocal language in an attempt to evade or diminish the psychological conflict produced by the discrepancy between the deceptive appealers' knowledge of their own guilt, and their appeals for help.

Deceptive appealers in the present studies also produced more gaze aversion than honest appealers, and this is in line with some previous research on high stakes deception, although, overall, the findings have been mixed. Gaze aversion has been associated with deception in a study investigating a police interview with a murderer (Vrij & Mann, 2001a), in a previous study on appeals (Wright Whelan, 2009), and in a meta-analysis of mostly low stake studies, when motivation to succeed in the lie was a moderating variable (DePaulo et al., 2003). However, in a different study of police suspect interviews, no significant effect of veracity on gaze aversion was found (Mann et al., 2002). As noted previously, gaze aversion as a cue to deception has strong theoretical underpinnings; gaze aversion has been related to cognitive load (Vrij & Mann, 2001a), emotional responses such as shame (DePaulo et al., 2003), and guilt (Ekman, 1992), and attempts by liars to be evasive and distance themselves from their communications (Zuckerman et al., 1981).

It may be that gaze aversion is more likely to emerge in some high stakes contexts than in others. For example, in the study of police suspect interviews in

which veracity was found to have no effect on gaze aversion, of the 16 crimes under investigation, nine were theft and two were arson, and the majority of the suspects were well known to police and had been interviewed on several occasions previously (Mann et al., 2002). It may be that an experienced criminal who is lying about a property crime, may not experience the same emotions (i.e. shame and guilt) that may be particularly relevant to somebody who is lying about a murder, especially the murder of their own relative. Furthermore, in the context of appeals, gaze aversion may be used as a form of distancing behaviour. As previously noted, honest appeals are likely to be a genuine attempt to communicate with the public and ask for help, and, therefore, honest appealers are more likely to behave in a way that facilitates this communication; for example, by using direct eye contact. The motivation of deceptive appealers is clearly not the same, as they do not actually want the public to assist in finding out what happened to their relative. Deceptive appealers may, therefore, produce behaviours that inhibit direct communication, such as gaze aversion. Clearly, further research is needed to establish under which specific high stakes circumstances, gaze aversion is more likely to be related to deception.

Deceptive appealers in the present studies also produced more head shakes than honest appealers, and this is in line with findings that increased headshaking related to deception in a study of police suspect interviews (Mann et al. 2006), and also in a study investigating cues to deception in criminal suspects' statements (Davis et al., 2005). As previously noted, the factor underlying the production of this cue to deception is unclear from the present studies; it may be a deliberately

exaggerated behaviour intended to convey an appearance of honesty (Davis et al., 2005), or it could be construed as ‘leakage’ of the negative emotions that the appellant is experiencing as a consequence of his or her deception (Ekman, 1992). The head shakes in the present study varied in magnitude from very small to quite large movements, and it could be that the former are more indicative of leakage, and the latter of exaggerated behaviour. This possible distinction may be worth following up in future research.

The finding of a trend for deceptive appellants to produce more speech errors than honest appellants offers some tentative support for previous results in this regard; for example, speech dysfluency and errors were found to increase in deceptive statements of criminal suspects to assistant district attorneys (Davis et al., 2005), and also in deceptive statements made in the police interview of a murder suspect (Vrij & Mann, 2001a). Again, there is strong theoretical rationale for this behaviour to relate to deception, as it is associated with both cognitive load (Vrij & Mann, 2001a), and anxiety (Davis et al., 2005).

On the whole, therefore, the present results appear to endorse the idea that some behaviours may discriminate between honesty and deception across a number of high stakes situations.

### **12.2.2. *Previously Unidentified Cues***

However, as hypothesised, a number of previously unidentified, and possibly context-specific cues, which derived from the investigation of cues used by accurate deception detectors (and also one noticed by the researcher), were also found to discriminate between honest and deceptive appeals. Clearly, as these cues have not been previously investigated, they require further validation; nevertheless, in as much as they were replicated with two different sets of observers, on two different sets of appeals, it would appear that they could potentially be important.

Most of the new cues identified in the present thesis concerned subjective impressions or observations; this is a type of measurement that has not been widely used in high stakes deception research (Koper & Sahlman, 1991; Mann & Vrij, 2006). Indeed, to the author's knowledge, the present thesis includes what seems to be first investigation of the personal reactions of observers to real life, high stakes liars.

Given the nature of the appeals context, it was hypothesised that emotion might be a key component in cues subjectively observed in this situation. And, accordingly, many of the subjective measures that discriminated between honest and deceptive appeals could be construed as directly relating to emotion; in particular, the concept of emotional authenticity (fake emotion, fake facial expression, performance, sad eyes, genuinely sad, genuine/heartfelt, urgency, personal and expressive voice, voice quivering with emotion). This suggests not only that cues

related to emotion may be useful in detecting deception in the context of appeals, but also that untrained observers may be able to discriminate between simulated and genuine emotion in this context. As such, the results could be considered to support not only the hypothesis that some aspects of genuine facial expression cannot be voluntarily simulated, or may be involuntarily expressed (Darwin, 1872; Ekman, 2001; Porter and ten Brinke, 2008), but also the idea that individuals in their everyday lives may be attuned to this.

The possible context specific nature of the emotional cues is perhaps worth particular emphasis. For example, a recent study of written statements by military personnel as witnesses or suspects in an investigation, found that verbal expressions of negative emotion were associated with deception (Fuller et al., 2013), and negative statements have also been related to deception generally (DePaulo et al., 2003). However, in the appeals examined in the present thesis, there was a trend for expressions of concern and pain (which could be considered to be negative emotions) to be more common in honest appeals than in deceptive appeals. Moreover, expressions of positive emotion were related to honesty, but when specifically referring to the relative. This again illustrates the point that use of contextual focus should be an essential part of the methodology of investigating high stakes lies (and, possibly low stakes lies).

Interestingly, as mentioned earlier, previous research using low stakes lies as stimulus materials has indicated that liars are sometimes judged to be less pleasant than truth-tellers (DePaulo et al., 2003), and that people feel less comfortable when



hearing a lie than when hearing a truth (Anderson et al., 1999). The present results appear to concur with these findings. Thus deceptive appealers were rated as being more creepy and unnatural than honest appealers, and observers disliked them more than honest appealers. Indeed, although all the subjective measures had very large effect sizes (all  $d > 0.80$ ), and gaze aversion was the most predictive cue (with 79% ‘hits’), the subjective cues creepy, and lack of sympathy for the appealers, had impressive ‘hit’ rates of 75%. Also, honest appealers were rated as more normal and plausible than deceptive appealers, and observers were more likely to report that they felt sorry for them, and felt their pain, more than they did for deceptive appealers.

The most popular explanation for the fact that liars are regarded as more unpleasant than truth-tellers is that the former are more likely to experience, and thereby manifest reactions to, unpleasant negative feelings of guilt and fear; i.e. because they manifest more evidence of moral conflicts, anxiety, negativity, and attempts at behavioural control, liars may come across as less pleasant than truth-tellers (DePaulo et al., 2003; Zuckerman et al., 1981). However, there may be other possible explanations of such findings which have not previously been considered. For example, as suggested on a number of occasions in this thesis, particularly in the context of appeals, such cues might also be construed as relating to norm violation. It could be the case that deceptive behaviour in itself violates certain presentational norms that usually go unnoticed; consequently, when it appears, it is judged ‘creepy’, ‘unnatural’, and generally unpleasant. In contrast, honest behaviour, which conforms to morally normative expectations, appears normal and plausible. However, perhaps

a more significant possibility, as yet unaddressed in the deception literature, is the possibility that, in the context of high stakes deception, the negative impressions of the observers were not a response to the act of deception per se, but were reactions to more global characteristics of people who choose to commit serious crimes and then lie about them. It can be noted that if this is the case, it would again illustrate the potential of the between subjects stimulus approach in uncovering cues to detection, as such differences would be unlikely to emerge in within subject comparisons. This possibility is explored more fully shortly.

Although the majority of the newly identified cues were subjective measures, three were verbal behaviours amenable to frequency counts: these were references to norms of emotion/behaviour, expressions of hope of finding the missing relative, and expressions of positive emotion towards the relative. Furthermore, honest appeals contained marginally significantly more expressions of concern for the relative or pain, and avoidance of the use of brutal language or detail. Again, these cues are not easily accommodated by traditional theoretical approaches to cues to deception, which have focused on emotions that may be associated with the act of deception: for example, fear, anxiety, guilt, shame and excitement (Ekman, 1992; Zuckerman et al., 1981). Such cues suggest that, in some high stakes contexts, such as that of making an appeal for a missing or murdered relative, the patterns of emotion may be more complex and context specific, than those found in studies concordant with the predominant Four Factor Model of detection deception (Zuckerman et. al, 1981), and may require different kinds of explanations. Again, this appears to endorse the utility

of a data-driven approach allowing cues to emerge that may point to new theoretical directions.

### ***12.2.3. Multiple-Cue Approach and Case by Case Analyses***

The second main aim of this thesis was to develop a possible multiple-cue approach to deceptive behaviours which discriminates between honesty and deception in the context of appeals, and which may be useful in an applied context.

As indicated in Chapter 10, the use of a range of cues to identify individual cases as honest or deceptive proved to be of considerable effectiveness, suggesting that a multiple-cue perspective using the kinds of cues identified here may be worth pursuing further. Notably, the only individual cue that was equivalent to the case by case analyses using frequency cues, in terms of the ability to predict veracity, was gaze aversion; however, it was not possible to measure this in every case, and this could be the case in a number of other applied contexts. Although the case by case analyses utilising subjective measures were reasonably successful in classifying appeals as honest or deceptive (69%), the case by case analyses using the cues measured by frequency counts were more successful (78%), though not significantly so. The latter included both verbal and nonverbal cues (gaze aversion, head shaking, equivocation, expressions of hope of finding the relative, expressions of positive emotion towards the relative, and references to norms of emotion/ behaviour),

suggesting that it is useful to assess cues across channels of behaviour. Importantly, using cues that can be measured with frequency counts is also potentially more useful in an applied context. The case by case analyses using cues measured by frequency counts also produced veracity classification accuracy that is more in line with previous research that has adopted this approach (Ekman et al., 1991; Leal et al., 2010), albeit not in a high stakes context.

This illustrates a general problem with using a restricted range of cues; i.e. in practice, it may not always be possible to identify a specific cue. Hence researchers such as ten Brinke and Porter (2012), and Vrij and Mann (2004) have argued a multiple-cue approach is more likely to be more successful in the field. However, perhaps most significantly, it appears that the kind of simple overall subjective judgement of veracity shown in Study 7 (Chapter 11) may be more useful as a predictor of veracity in high stakes contexts, than even a multiple-cue approach.

### **12.3. Theoretical Development**

The third main aim of the present thesis was to develop a theoretical rationale for cues that may be context-specific and may emerge only in high stakes contexts, and cannot readily be accommodated by existing theoretical approaches.

As has already been emphasized, given the variety of cues to deception that were identified in the present thesis, it is difficult to explain them in terms of any existing, single, overarching, theoretical perspective; the dominant approaches to the production of cues to deception do not fully account for a number of the context-specific cues identified in the present thesis, although some are more readily accommodated by existing theoretical viewpoints than others. For example, whilst it may be possible to explain some cues (gaze aversion, speech errors, equivocation, head shaking), in terms of general factors such as cognitive load, affective reaction, arousal, and impression management, in addition to these more familiar cues, the present studies also generated some more novel cues that are less obviously explicable in these ways, and might be more specifically related to the context of making an appeal.

For instance, references to norms of emotion/behaviour were significantly more prevalent in honest appeals than in deceptive appeals, and honest appeals contained significantly more expressions of hope of finding the missing relative alive than deceptive appeals, in cases in which a body had not been found. Honest appeals were also significantly more likely to include an expression of positive emotion towards the relative than deceptive appeals. Furthermore, honest appeals contained marginally significantly more expressions of concern or pain, and avoidance of the use of brutal language. As noted previously, these findings would appear to endorse the view that, although theoretical approaches to cues to deception have traditionally focused on emotions that specifically relate to the act of deception, for example, fear,

anxiety, guilt, shame and excitement (Ekman, 1992; Zuckerman et al., 1981), in high stakes contexts, such as that of making an appeal for a missing or murdered relative, the patterns of emotion may be more complex and context specific, requiring a different kind of explanation. Two possibilities are as follows.

### **12.3.1 *The Social Interactionist (SI) Account***

Broadly speaking, these additional, more context specific cues might be better accommodated by what we could term a ‘social interactionist’ (SI) perspective. As Salfatti (2003) says, ‘Homicide grows out of a transaction between individuals. This transaction is a product of the individuals and their relationship’ (p.499). That is, within the context of the kinds of cases used here, where there are offenders and victims, some cues to deception reflect the nature of the social relationships between the offender and the victim, and corresponding attitudes. So, for example, whilst it might be expected that somebody making an honest appeal would be more likely to exhibit genuine sadness about the loss of his or her relative, somebody who has recently killed his or her relative may not experience the same sense of loss (indeed, the act of killing the victim would imply that they might even be pleased or relieved). Similarly, whilst it might be expected that somebody making an honest appeal would be more likely to express positive emotion towards his or her relative, somebody who has recently killed his or her relative may lack the same positive emotion

towards the relative (indeed, the act of killing the victim could be construed as an obvious manifestation of this).

Along the same lines, one might expect that someone making an honest appeal would produce more expressions of concern or pain about the disappearance or death of the relative, than somebody who has recently killed the relative (and was presumably motivated to do so). Moreover, one might expect honest appealers to display more urgency in their attempt to find the relative, or to find out who killed their relative, than deceptive appealers who clearly do not have the same motivation regarding their relative. Additionally, honest appealers were rated as having a voice quivering with emotion more than deceptive appealers; the traditional arousal approach would predict that liars would exhibit more emotional stress in their voices than truth-tellers, but according to the social interactionist approach, one might expect more emotion in the voice of somebody who is genuinely concerned or grieving for their relative, than somebody who has killed their relative and so may not be experiencing these emotions. Also, whereas an honest appealers may be averse to expressing any brutality towards their relative, a deceptive appealers may not be so reluctant.

In cases in which the relative was not publicly known to be dead, honest appealers were also found to produce more expressions of hope of finding the missing relative, than deceptive appealers. From a social interactionist perspective, this might simply be a cognitively driven behaviour that reflects an obvious change in the status of the victim within their interaction; i.e. the deceptive appealers (except

in one case where the relative was kidnapped but still alive), knew that their relative was already dead and would, therefore, have no spontaneous reason for expressing such hope. Furthermore, honest appeals contained more references to norms of emotion/behaviour (including violations of norms) than deceptive appeals; this may, in part, also be a cognitively driven behaviour reflecting the nature of the interaction between offender and victim. When honest appealers ask, ‘how could somebody do this?’ or, ‘why would anybody want to hurt her?’ they are presumably asking a genuine question to which they do not know the answer. A deceptive appealers, however, knows from personal experience how and why somebody could kill their relative. Although an additional possibility is that, in an attempt to mitigate cognitive dissonance, deceptive appealers may be less likely to refer to norms of emotion and behaviours which they know that they have violated.

Importantly, this SI approach is useful in explaining newly identified, context specific cues in a different high stakes situation. Several such cues were identified in Harpster et al.’s (2009) study of 911 homicide calls, in which some of the callers were involved in the homicide which they were reporting. For example, as in appeals examined in the present thesis, honest callers (i.e. those who were not involved in the homicide they were reporting), expressed more urgency and were more demanding in the requests for assistance, than deceptive callers. As in appeals, it may be that this reflects the nature of the relationship between the offender and the victim, i.e. deceptive callers do not have the same motivation to help the victim as honest callers. Similarly, honest callers were more likely to make a plea for help for the



victim, whereas deceptive callers were more likely to make a plea for help for themselves; again, one might expect somebody who has killed a victim to be less likely to plea for help for that victim. Deceptive callers were also more likely than honest callers to insult the victim, and to describe the death of the victim in terms of a problem, and again, according to the SI approach, these behaviours could be regarded as reflecting the nature of the social interaction between the offender and the victim: one might expect a caller who has just killed a victim to be more likely to insult them, or regard their death as a problem, than a caller who is genuinely trying to help the victim. In a similar vein to this, a recent study reported that in a significant number of cases where offenders had made false allegations of child abductions, the victim was described by the offender as ‘difficult’ or ‘different’ (Canning, Hiltz & Muirhead, 2011). Furthermore, in Harpster et al.’s study, honest callers were less likely than deceptive callers to accept that the victim was dead, and this behaviour is similar to honest appealers, who were more likely than deceptive appealers to express hope of finding their relative alive.

The SI approach also potentially accommodates some of the more context-specific findings reported by ten Brinke and Porter (2012), in their analysis of appeals. The researchers found that deceptive appealers were more likely than honest appealers to produce partial facial expressions of disgust and happiness, and less likely than honest appealers to produce facial expressions of sadness/distress. According to the SI approach, one would expect appealers who had actually killed their relative to be more likely than honest appealers to leak expressions of disgust

when talking about their relative, or to leak expressions of happiness when talking about the absence of their relative. Similarly, as with the findings of the present thesis, one would expect honest appealers to exhibit expressions of sadness/distress more than deceptive appealers, who have already killed their relative and were presumably motivated to do so.

### **12.3.2. *The Individual Behavioural Profile (IBP) Account***

A second, and related, way of interpreting some of the newly identified, context specific cues in the present thesis, is in terms of what we can call the ‘individual behavioural profile’ (IBP) account of deception. This has also yet to be systematically applied to studies of deception. According to this view, some cues may reflect general characteristics associated with the kinds of individuals who commit the crimes that the deceptive appealers have committed in a particular context (and who subsequently lie about it for obvious reasons), rather than their acts of deception per se. That is, the cues may form part of the behavioural profile of individuals who engage in particular aberrant acts. In the present cases, and with most other high stakes deception contexts, honest and deceptive people differed not only in terms of whether they were lying, but also whether they had committed serious criminal acts.

It is notable, for example, that, as with lying itself, showing fewer signs of genuine emotion, fewer signs of positive emotion towards the victim, and less concern towards the victim, a failure to use references to appropriate norms of emotion/behaviour, and the use of brutal language, could all be construed as consistent with the stereotype of a psychopath (Davies & Feldman, 1981). In particular, the failure to refer to norms of emotion/behaviour is interesting, as norms imply the assumption of a shared understanding which, in relation to behaviour and emotion, a psychopath may lack. Also, being disliked, being regarded as unnatural, creepy, not normal, and not worthy of sympathy, might reflect characteristics of people who murder relatives, rather than characteristics related to the act of deception per se. In support of this proposition is the finding that there was no relationship between the percentage of correct credibility judgements, and the length of the appeal; this implies that observers may have been making accurate, rapid, global judgements, rather than relying on careful analysis of lengthy sections of behaviour. It may be, then, that observers were identifying an aberrant 'type' of person, rather than specific cues to deception; this also fits with the finding that, for the deceptive appeals, overall correct veracity judgements were not related to 'hits' on the cues measured by frequency counts, but were related to 'hits' on the cues measured by subjective observer ratings (which mainly comprise global impressions). It may be that observers have an expectation of what 'normal' behaviour encompasses in this context, and violations of this expectation imply that the appellant is, in fact, not 'normal', but is aberrant.

As emphasised previously, behaviours of this kind will obviously not emerge as discriminatory in standard low stakes laboratory situations where ‘normal’ participants are randomly assigned to lying and truthful conditions. Moreover, because they may not necessarily relate to the act of deception per se, they may not be apparent in within-subjects comparisons of truthful and deceptive conditions, even in high stakes situations. Consequently, whilst, at a practical level, real-life high stakes situations might potentially provide a useful source of cues for making probabilistic estimates about who is lying and who is not, they may, nevertheless, pose some methodological problems for researchers attempting to determine cues to deception per se, as distinct from cues that identify more global characteristics of offenders (who, through their actions, end up lying).

#### **12.4. Observer Accuracy in Detecting Deception**

The fourth main aim of the present thesis was to investigate the effect of using ecologically valid stimulus materials on the accuracy of observers (both police and non-police) in detecting deception, on truth bias, on accuracy-confidence relationships, and the utility of consensus judgements as a predictor of veracity. It was hypothesised that observers would be able to judge veracity in the appeals with greater accuracy than is usual in low stakes deception studies, as the detection of

deception necessarily depends on discernible differences between honest and deceptive communication, and the factors which underlie the production of cues to deception are likely to be exacerbated in high stakes situations. The hypothesis was supported; reports from mainly low stakes studies have previously indicated observer accuracy in detecting deception from simple judgements of veracity to be in the region of 54% (Bond & DePaulo, 2006); i.e. only slightly above chance. In contrast, the overall accuracy of the sample of observers in Study 7 was 71%; which was significantly above chance.

#### **12.4.1. *General Accuracy***

Given that the importance of the ability to detect deception is fundamental to police work, the police were an obvious group to test as a possible source of individual differences in observer accuracy in detecting deception.

Police participants and participants from the general public, achieved accuracy rates well above those expected from studies using low stakes stimulus materials, at 72% and 68% respectively (it can also be noted that in studies 2 and 4, in which a lower number of non-police participants and/or stimulus materials were used, similar accuracy rates of 75% and 67% were found). Furthermore, both groups also achieved accuracy rates comparable to those achieved in a small body of research investigating police accuracy in detecting deception when viewing police

suspect interviews, of 68%, 65%, 69%, 64% and 72% (Mann & Vrij, 2006; Mann et al., 2004, 2006; Vrij & Mann, 2001b; Vrij et al., 2006). This has two important implications. First, these findings suggest that, as expected, police participants are able to achieve high accuracy levels in a high stakes context other than that of police suspect interviews (although the context of appeals could still be regarded as forensic). However, second, to the author's knowledge, the present studies are the first in which the accuracy of participants from the general public, when viewing real life, high stakes lies, has been investigated. The findings strongly suggest that the common proposal that people are poor at detecting deception (for example, Bond & DePaulo, 2006), may be an artefact of the methodology utilised; that is, people may be poor at detecting low stakes lies, in which cues to deception are weak and difficult to discern, but they are able to discriminate between honesty and deception in the high stakes context of appeals with relatively high accuracy.

As previously noted, there has been lively debate regarding the existence of experts in deception detection (see, for example, Bond & Uysal, 2007; O'Sullivan, 2007; Shaw et al., 2013). However, it could be the case that with more ecologically valid stimuli, the ability to detect deception may not be as uncommon as some researchers have suggested. For example, O'Sullivan and Ekman (2004) tested thousands of individuals and identified only a handful who are able to detect deception above 80% on two different tasks (both low stakes). However, previous preliminary research using very limited samples of high stakes stimulus materials has indicated that this ability may not be so unusual; a number of police officers have

generally achieved accuracy rates of 80% and higher in several previous studies (Mann et al., 2006; Vrij & Mann, 2001a; Vrij & Mann, 2001b). These findings were replicated in the present thesis using a much larger sample of stimulus materials (36 appeals). Indeed, in Study 7, 15 individuals achieved 80%+ accuracy. Importantly also, this included four individuals from the general public; as noted above, the accuracy of participants from the public when viewing high stakes lies has not previously been investigated.

#### ***12.4.2. Police vs Non-Police: Accuracy, and Truth and Deception Bias***

As hypothesised, police participants were more accurate overall than participants from the public, and this is an important finding because, as far as the author is aware, there have been no previous investigations comparing police accuracy with public accuracy, when judging real life, high stakes lies. Significantly, however, although police participants were more successful than the public at correctly identifying deceptive appeals, there was no difference in accuracy between police and non-police participants in correctly identifying honest appeals; thus it was the police officers' ability to correctly identify deceptive appeals that accounted for their higher overall accuracy rates. Nevertheless, the proposal that police participants may achieve higher accuracy than participants from the public because they are exposed

to more real life, high stakes lies, was apparently not supported in that years of experience as a police officer was not related to accuracy.

However, some of the findings regarding truth bias may be helpful in offering some explanation for this. It was hypothesised that, because of the high stakes nature of the stimulus materials, cues to deception were likely to be more discernible, and, therefore, truth bias would be eliminated. This hypothesis was not supported in relation to participants from the public, who displayed a truth bias despite viewing real life, high stakes lies. However, as expected, police officers actually showed some evidence of a deception bias. One possible explanation for this is that, perhaps because of their training, and even limited experience in the field, in terms of Signal Detection Terms, Police officers are encouraged to adopt a more lax criterion (beta shift) for the classification of deception, and a more cautious criterion for the classification of honesty, than the public. This would result in them paying more initial attention to cues for dishonesty than honesty. This, coupled with any benefits of more experience with people who (because of other evidence) are known to be lying, may give the police an advantage over the public that plateaus with a certain amount of experience (i.e. it will not necessarily improve with years of experience). It can be noted that in the sample used here, the minimum experience was two years. If this is the case, this suggests that only limited exposure to a range of individuals who are lying may be sufficient to boost the ability to detect deception.



#### **12.4.3. *Relationships Between Accuracy and Confidence***

It was hypothesised that, due to the use of real life, high stakes lies as stimulus materials, in which cues to deception are likely to be more discernible, observers would be more confident in their accurate veracity judgements, than in their inaccurate veracity judgements. This hypothesis was supported. This is potentially an important finding given that previous, mostly low stakes, studies have found that people are no more confident when making accurate veracity decisions than when making inaccurate veracity decisions (DePaulo et al., 1997). Again, the basic premise of the present thesis, that investigation of high stakes deception will produce different results to investigation of low stakes deception, is endorsed. Previous findings have used only police participants, and have produced conflicting results (Mann et al., 2004, 2006; Vrij & Mann, 2001a). However, as before, these studies used a very limited range of stimulus materials, thus restricting both the potential for variance in the data, and the robustness of the results. The present results, therefore, could be considered to endorse the view suggested by other research on witness accuracy and confidence (Kebbell et al. 1996; Wheatcroft et al., 2004), that, with a wide range of stimuli that include easy or ‘obvious’ items, the more confident people are that they are correct, the more likely they are to actually be correct. To the author’s knowledge, there have been no previous investigations of the accuracy-confidence relationship in participants from the public, when judging high stakes

lies, so the present findings may be particularly important in this respect, and worthy of further investigation.

Interestingly also, participants were more confident when correctly identifying deceptive appeals, than when correctly identifying honest appeals, implying that deceptive appeals may be more ‘obvious’ to observers, than honest appeals. As discussed previously, this could be due to a number of factors. For example, cues to deception tend to be more numerous and are, perhaps, more prominent than cues to honesty; thereby making them more salient or ‘easier’ to recognise. Also, in this context, again, one particular factor that might contribute to their prominence is norm violation; they are more prominent because they violate the norms of ‘usual’ acceptable behaviour in social contexts.

#### ***12.4.4. Consensus Judgements as a Predictor of Veracity***

As expected, observer consensus on credibility predicted veracity; hence when observers agreed on a simple overall veracity judgment, they tended to be correct. This finding was important for two reasons. First, to the author’s knowledge, observer consensus as a predictor of veracity has not previously been investigated, and the results from the present thesis suggest that it is an area that warrants further research. Second, simple consensus judgments were very highly predictive of veracity, correctly classifying 33 out of 36 appeals (92%); this was a higher ‘hit’ rate

than any individual cue, or any combination of groups of cues. Clearly, as the utility of consensus has not previously been investigated, replication of the present findings is needed; nonetheless, in the context of appeals at least, it seems that senders display sufficient deceptive and honest behaviours for the majority of observers to make correct veracity judgements about their source. Moreover, as previously noted, the present results suggest that it may be as, if not more effective to use aggregated judgements, rather than investigating differences in behaviour between liars and truth-tellers, to predict veracity.

#### **12.4.5. *Relationships Between Consensus and Cues***

The hypothesis that appealers who were accurately identified as honest or deceptive by observer consensus, would also exhibit the cues identified in Studies 4, 5 and 6 as reliably discriminating between honesty and deception, was only partially supported. There was no significant relationship between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ on the verbal and nonverbal cues measured by frequency counts in Studies 5 and 6. However, in the deceptive appeals, there was a strong correlation between the percentage of correct veracity judgements on each appeal, and the percentage of ‘hits’ on the subjective cues rated by observers in Study 4. This finding again suggests that deceptive behaviour may violate norms which go unnoticed when conformed to, but which are noticed when

violated. It also suggests that, as consensus was not related to cues measured by frequency counts, observers may have been relying on more global judgements such as those related to the SI and IBP accounts detailed earlier (for example, ‘creepy’ or ‘unnatural’), rather than specific behaviours, to make their accurate veracity judgements of deceptive appealers. Another, or additional possibility, of course, is that, notwithstanding the fact that the cues generated by the accurate observers in Study 1 were relatively successful in predicting veracity, observers in general may sometimes be poor at identifying and recognising the cues they used to come to their decisions; indeed, as evidenced in Nisbett and Wilson’s (1977) seminal paper in the subject, people can be notoriously poor at accurately recognising and reporting their own mental processes.

## **12.5 Methodological Limitations**

Despite what could be considered to be the promising nature of many of the findings from the studies in this thesis, the studies themselves obviously suffer from a number of important methodological limitations which constrain their generalisability. It can be noted that the fact that the stimulus materials concerned only the context of appeals is not in itself a methodological limitation as this was the focus of the thesis; i.e. the primary intention was actually to identify cues that may be contextually

dependent. Nevertheless, even allowing for this, there remain a number of potentially problematic issues.

As the research in the present thesis was based on field data, key experimental elements are missing (for example, randomisation and manipulation). However, it was considered that this limitation was outweighed by the need for ecologically valid data; hundreds of laboratory-based experiments on deception, in which variables are controlled and manipulated, have already been conducted, but realistically replicating a high stakes environment is likely to be impossible in the laboratory. The aim of the present thesis was not to replicate these kinds of studies, but to contribute to the small body of research on high stakes lies, in order to expand our understanding of real world deceptive behaviour. Therefore, despite these limitations, it was considered that the advantages of the stimulus materials hugely outweighed the drawbacks.

Nevertheless, notwithstanding these arguments, there are a number of other limitations associated with the stimulus materials that need addressing in further research. For example, all the appeals examined in the present thesis were made in the United States, the United Kingdom, Canada and New Zealand. Hence the Western bias of the sample must be acknowledged. For the purposes of the present thesis, it was necessary to use English speakers to investigate verbal behaviours and also observer accuracy. It is certainly possible that some behaviours may not generalise across cultures. For example, gaze behaviour, smiling and expressions of emotion have been shown to differ between Eastern and Western cultures (Argyle,

Henderson, Bond, Iizuka & Contarello, 1986; Jack, Caldara & Schyns, 2012). A related issue here, of course, concerns the sample size used per se. Although every effort was made to gain as large a sample of appeals as possible, even the final sample was limited. Hence the studies need repeating on a larger and more representative sample of stimulus materials before they can truly be considered to be robust.

Another problem inherent in research using real life lies is the issue of determining ground truth; the fact that an individual may be found guilty or not guilty by a court may not in itself be a reliable indicator that he or she committed the offence in question. In the present thesis considerable efforts were made to address this concern, and cases were included only if they satisfied the criteria used in other published research in this area by Vrij and Mann (2001b), and ten Brinke and Porter (2012); thus it is unlikely that a significant number of cases were incorrectly classified as honest or deceptive. However, it should be born in mind this remains an issue that can lead to an underestimation of the cues available and observers' accuracy in identifying them.

Another problem relating particularly to the stimuli used here is that the resolution of the footage used was variable, such that some cues could not be investigated at all or were only available for a limited number of cases (for example, blink rate and pupil dilation). The scope of the image also varied in the appeals, such that in some cases just the head was shown, limiting the investigation of further nonverbal behaviours (for example, illustrators and self-manipulators). This suggests

that, in future research, attention should be paid to providing higher quality video footage. It could be that some cues were missed or underestimated because of this problem. This also raises the issue of whether video footage collected in this way is necessarily the best source of information for eliciting cues to deception in the context of appeals. For example, it would obviously make more sense for the investigators to actually attend such briefings to control the filming and audio recording that takes place; though presumably, with the resources available to most investigators, this is unlikely to be practicable.

However, perhaps the most obvious limitation with regards to the present study was the size of the participant samples. The qualitative data were based on only two observers, and the subjective ratings were done by only five raters. Nevertheless, it should be emphasized that, in terms of both raters and the number of stimulus materials, the present studies are considerably superior to most previous studies on this topic. For example, it is usual in deception research to use as few as two or three people to rate behaviours (Davis et al., 2005; Harpster et al., 2009; Koper & Sahlman, 1991; Mann et al., 2002; ten Brinke & Porter, 2012; Villar et al., 2011; Vrij & Mann, 2001a), and a body of research has been based on police interviews of 16 suspects (Mann, et al., 2002, 2006; Mann & Vrij, 2006), whilst some studies have included stimulus materials produced by only one ‘sender’ (Villar et al., 2011; Vrij & Mann, 2001a). Furthermore, it could be argued that if the cues reliably emerge with such a small number of raters, this attests to their salience and robustness. However, by the same token it could be argued that there is still the potential for some of the

cues identified to be ‘false alarms’ and for some cues to have been missed. In this respect, therefore, the present results should be treated as provisional and awaiting confirmation on larger samples.

In relation to the analyses conducted, regression analysis would clearly have been more suitable for multiple-cue prediction; however this was not possible for two reasons. First, in the group of cues measured by frequency counts, two of the predictor variables (gaze aversion, and hope of finding the missing relative) were not available in all cases. Second, in the group of cues measured by subjective observer ratings, there were too few cases per predictor variable, resulting in a danger of over-fitting the model to the data. This relates to the problem of the limited sample size, just discussed. There are only a certain number of these types of cases in existence, and of those, only a certain number of appeals are available to download freely (although the available data is increasing every year). However, in any case, for applied purposes, a simple checklist of behaviours, which can be used without conducting statistical analyses, would seem to be most appropriate.

Also, in investigating the cues to deception, one of the main methods used was to count frequencies of behaviour. This is the standard approach in deception research, but is limited in that behaviours are not considered in relation to each other. Informal, post-study interviews with a number of individuals identified as highly accurate in detecting deception in Study 7, suggested the importance of considering behaviours in relation to each other. A few limited attempts were made to explore this possibility (for example, including the cue ‘smile appropriate to speech content’



and ‘calmness appropriate to speech content’ in Study 2), but without success. This type of approach, with concurrent attention to multiple channels of behaviour, has been advocated by Porter and ten Brinke (2010), who have also suggested that some manifestations of deception are idiosyncratic and difficult to address in quantitative research. Possible examples of this in the present thesis include comments by the observers in Study 1 that deceptive appealers tended to make assumptions about their missing relative, or that there were indications of a dispute with the missing relative. Such behaviours may occur too rarely to be tested systematically, but, when they occur, they may contribute to a holistic picture which enables the observer to accurately detect deception.

## **12.6 Suggestions for Future Research**

As the body of research investigating real world, high stakes deception is small, there are as yet many unanswered questions. For reasons previously suggested, at this stage the present results can only be considered provisional or preliminary, but it is hoped they can contribute to our understanding of deceptive behaviour, and also point to possible areas in which further investigation might be useful.

In terms of future research, an obvious first point to make is that all of the present studies need replicating with different and hopefully larger samples and

ranges of stimulus materials and participants. It would be particularly interesting to know whether, for example, they generalise across cultures and to different high stakes situations.

Of the cues found to discriminate between honest and deceptive appeals which have also been found to discriminate between honesty and deception in other high stakes contexts, two in particular would benefit from further investigation. Considering the inconsistent findings regarding gaze aversion in the high stakes deception literature, research exploring the specific circumstances under which gaze aversion is related to deception, and the factors likely to be underlying the production of this behaviour, would be useful. Additionally, the factors underlying the production of head shaking are unclear and warrant further investigation; it may be that large movements are a result of deliberate behaviour and are indicative of attempts at impression management, whereas small movements may indicate leakage of negative emotion.

A large number of novel cues were identified in the present thesis, and as with any novel findings, these, in particular, need validating in further studies. Nevertheless, the findings of the present thesis suggest that considering a specific contextual focus, and utilising a fine-grained, data-driven approach, may result in the emergence of context specific cues, and the development of a multiple-cue approach to behaviours predictive of deception in a specific context; this could be developed in other specific, high stakes contexts.

And again, it would also be useful to investigate which, if any, of the newly identified cues have utility in other high stakes contexts. As has already been mentioned, there appears to be some overlap between the cues found to discriminate between honest and deceptive appealers, and the cues found to discriminate between honest and deceptive 911 callers. It could be the case, therefore, that some cues have utility beyond the specific context of appeals.

Although multiple channels of behaviour were investigated in the present thesis, it was beyond the scope of the research to investigate discrepancies between channels of behaviour. This is an approach that is very rarely adopted in deception research, probably due to the difficulties in operationalization, and yet it is an important part of the strategies used by individuals accurate at detecting deception (for example, O'Sullivan & Ekman, 2004), and is an issue that future research could attempt to address. Kosslyn (as cited in Henig, 2006) has described deception as 'a huge multidimensional space in which every combination of things matters'; this view has been echoed by Porter and ten Brinke (2012), who suggest that combinations, synchrony, asynchrony, and context of behaviours all matter, and that a nuanced, dynamic approach is needed for advances in deception detection in the field.

Research reporting observer accuracy in detecting high stakes deception is scant, and although the present findings contribute to this small body of research, there is wide scope for further investigation of this area. For example, replication is needed of the findings regarding the high accuracy of non-police observers, and also

of the relationship between accuracy and confidence in this group. Also, replication is needed of the finding that police observers are better able to detect deception than non-police observers, and of the factors that may underlie this difference. The finding that observer consensus was a very strong predictor of veracity also clearly warrants further investigation, not only for academic reasons, as an interesting and potentially important psychological phenomenon, but for practical reasons as noted in the next section.

Finally, given that the findings of the present thesis resulted in the development of new theoretical approaches, there is plenty of scope for future research to investigate the utility of the new theoretical accounts in other high stakes contexts. It could be the case that the social interactionist and individual behavioural profile approaches may be particularly relevant to contexts which involve violent crimes in which there are offenders and victims.

## **12.7. Practical Applications and Implications**

One of the central themes of the present thesis was that, for any results to have practical significance, it is particularly important to use ecologically valid stimulus materials in deception research. It was, therefore, hoped that, by using more

ecologically valid stimulus materials, the studies in the present thesis, if replicated and developed, might have implications beyond the laboratory.

Appeals are routinely watched by police officers who are investigating the case of the missing or murdered relative, and who may well be forming credibility judgements about the appellant during the appeal; it is important, therefore, that there exist some criteria to inform these credibility judgements. Notwithstanding the caveats already mentioned, one particular finding that could potentially have practical implications here is that, in this kind of high stakes context, people are able to make accurate subjective global credibility judgements, and that a consensus of these judgements is particularly useful in predicting veracity. Indeed in the research presented here, this kind of consensus judgement was potentially a more accurate predictor of veracity than even a multiple-cue approach. On first consideration, in terms of efficiency, this could be considered promising as an approach in actual investigations; i.e. to make a fairly accurate prediction of the veracity of a sender's statement, all one needs to do is gather together a group of uninformed observers and ask them to make a global subjective judgment as to whether or not the sender is lying. Indeed, the present research appears to contradict the view, that for accurate detection of deception, observers need to be highly trained in this particular field; though there is some indication that police officers may fare better.

However, in practical terms, there are several issues to consider. It has been extensively demonstrated in the deception literature that in low stakes situations, simple observer judgements are not effective in predicting veracity (Bond &

DePaulo, 2006). Consequently, as low to high stakes contexts can be construed as existing on a continuum, it may not always be possible to determine at what point consensus judgements become more valuable than assessing cues to deception; thus in practical terms it may be beneficial to use both a multiple-cue approach and an investigation of consensus judgements. Also, as previously discussed, the independence of individual judgements may be a crucial factor in the accuracy of overall consensus in predicting veracity, and in many real life contexts, this independence may be difficult to ensure.

With regard to results pertaining to particular cues, hopefully the present findings may be useful in directing attention to some of the behaviours amenable to frequency counts, that best discriminate between honest and deceptive appealers: potentially the most useful candidates in this respect appear to be gaze aversion, head shaking, equivocation, expressions of hope of finding the relative, expressions of positive emotion towards the relative, and references to norms of emotion or behaviour. Moreover, the present research suggests that, to use these, one does not need to use highly trained specialists, or apply some kind of computer generated algorithm, to make sense of such data; a simple checklist may suffice.

It can also be noted that some of the findings presented here may be relevant to other contexts (for example, gaze aversion, equivocation, and head shaking). Though, as suggested in the previous section, perhaps particularly significant are those that could be identified with the social interactionist and individual behavioural accounts of deceptive behaviour. It could be the case that similar cues may emerge in

a variety of other contexts, but particularly in violent crimes in which there are offenders and victims.

## **12.8. Conclusions**

The basic premise of the research presented in the present thesis, was that the use of real life examples of honest and deceptive behaviour in a high stakes context as stimulus materials, would result in findings different to those typically produced in low stakes deception research. This premise was supported; a large number of behaviours were found to differ between honest and deceptive individuals in the context of appeals (all with effect sizes much larger than is common in deception research), and observers were able to detect deception with relatively high accuracy. It is hoped that these basic findings will reinforce the importance of ecological validity in deception research. Additionally, the research presented here has other implications that it is hoped will be of interest to researchers in the field of deception; that it may be important to consider context, a factor which is often overlooked in deception research; that the ability to detect deception in high stakes contexts may not be as unusual as previously thought, and that investigating cues used by individuals able to do this may be a fruitful area for research; and that simple consensus judgements may be the most effective method of classifying individuals as

honest or deceptive. Furthermore, it is hoped that the development of two new theoretical approaches to the production of deceptive behaviours may point to previously unexplored areas of possible research.



## References

- Adams, S.H. & Jarvis, J.P. (2006). Indicators of veracity and deception: an analysis of written statements made to police. *The International Journal of Speech, Language and the Law*, 13 (1), p.1-22
- Anderson, D.E., DePaulo, B.M., Ansfield, M.E., Tickle, J.J. & Green, E. (1999). Beliefs about cues to deception: Mindless stereotypes or untapped wisdom? *Journal of Nonverbal Behaviour*, 23, 67-89
- Argyle, M., Henderson, M., Bond, M., Iizuka, Y. & Contarello, A. (1986). Cross-cultural variations in relationship rules. *International Journal of Psychology*, 21 (1-4), 287-315. doi: 10.1080/00207598608247591
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91 (2), 276-292.
- Blair, J. P., Levine, T. R. & Shaw, A. S (2010). Content in context improves deception detection accuracy. *Human Communication Research*, 36, 423-442. doi:10.1111/j.1468-2958.2010.01382.x
- Blix, M., Wadejford, J., Wienecke, U., & Adahl, M. (2001). How good is the forecasting performance of major institutions? *Economic Review*, 3, 451–476.
- Bond, C. F. & DePaulo, B. M. (2006). Accuracy of deception judgements. *Personality and Social Psychology Review*, 10 (3), 214-234. doi: 10.1207/s15327957pspr1003\_2

- Bond, C. F. & Uysal, A. (2007). On lie detection “wizards”. *Law and Human Behavior*, 31 (1), 109-115. doi: 10.1007/s10979-006-9016-1
- Bond, G. D. (2008). Deception detection expertise. *Law & Human Behavior*, 32, 339-351. doi:10.1007/s10979-007-9110-z
- Buller, D. B. & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, 6 (3), 203-242
- Burgoon, J.K., Blair, P. & Strom, R.E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research*, 34, 572-599. doi:10.1111/j.1468-2958.2008.00333.x
- Canning, K. E., Hilts, M. A. & Muirhead, Y. E. (2011). False allegation of child abduction. *Journal of Forensic Sciences*, 56 (3), 794-802. doi: 10.1111/j.1556-4029.2011.01715.x
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum
- Cohen, D., Beattie, G. & Shovelton, H. (2010). Nonverbal indicators of deception: How iconic gestures reveal thoughts that cannot be suppressed. *Semiotica*, 182, 133-174. doi:10.1515/semi.2010.055
- Deceive. (2010). In *Collins English Dictionary* (30<sup>th</sup> ed.). London: Collins UK
- Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L.B., Parvizi, J. & Hichwa, R.D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, 3 (10), 1049-1056

- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. Chicago: University of Chicago Press.
- Davis, M., Markus, K.A., Walters, S.B., Vorus, N. & Connors, B. (2005). Behavioral cues to deception vs. topic incriminating potential in criminal confessions. *Law and Human Behavior*, 29 (6), 683-704. doi: 10.1007/s10979-005-7370-z
- Davies, W. & Feldman, P. (1981). The diagnosis of psychopathy by forensic specialists. *British Journal of Psychiatry*, 138, 329-331
- DePaulo, B. M. & Bond Jr, C. F. (2012). Beyond accuracy: Bigger, broader ways to think about deceit. *Journal of Applied Research in Memory and Cognition*, 1 (2), 120-121. doi:10.1016/j.jarmac.2012.04.010
- DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J. & Muhlenbruck, L. (1997). The Accuracy-Confidence Correlation in the Detection of Deception. *Personality and Social Psychology Review*, 1 (4), 346-357. doi: 10.1207/s15327957pspr0104\_5
- DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, 16, 249–267.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70, 979–995

- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. and Cooper, H. (2003). Cues to Deception. *Psychological Bulletin*, 129 (1), 74-118
- DePaulo, B. M. & Morris, W. L. (2004). Discerning lies from truths: behavioural cues to deception and the indirect pathway of intuition. In Granhag, P. A. & Stromwall, L. A. (Eds), *The Detection of Deception in Forensic Contexts* (pp.15-40). Cambridge: Cambridge University Press
- Dilmon, R. (2009). Between thinking and speaking – Linguistic tools for detecting a fabrication. *Journal of Pragmatics*, 41, 1152-1170
- Dolcos, F., LaBar, K.S. & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of The National Academy of Sciences of the United States of America*, 102 (7), 2626-2631
- Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, marriage, and politics* (3rd ed.). New York: Norton
- Ekman, P. & Friesen, W.V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, 32 (1), 88-106
- Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, 46, 913–920.
- Ekman, P., O'Sullivan, M. & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, 10 (3), 263-266

- Ekman, P., O'Sullivan, M., Friesen, W.V., Scherer, K. R. (1991). Invited article: Face, voice and body in detecting deceit. *Journal of Nonverbal Behavior*, 15 (2), 125-135
- Enos, F., Benus, S., Cautin, R.L., Graciarena, M., Hirschberg, J. & Schriberg, E. (2006, September). Personality Factors in Human Deception Detection: Comparing Human to Machine Performance. Paper presented at the International Conference on Spoken Language Processing, Pittsburgh.
- Eriksson, A. & Lacerda, F. (2007). Charlatanry in forensic speech science: A problem to be taken seriously. *International Journal of Speech, Language and the Law*, 14 (2), 169-193.
- 'Eyes for Lies' (2007). Looking at Matthew Gretz (Kira Simonian). Retrieved online 1 November 2009, from <http://blog.eyesforlies.com/2007/07/looking-at-matthew-gretz-kira-simonian.html>
- 'Eyes for Lies' (2008a). Penny Boudreau. Retrieved online 1 November 2009, from <http://blog.eyesforlies.com/2008/02/penny-boudreau.html>
- 'Eyes for Lies' (2008b). William Walsh. Retrieved online 1 November 2009, from <http://blog.eyesforlies.com/2008/10/william-walsh.html>
- Feeley, T. H. & deTurck, M. A. (1998). The behavioral correlates of sanctioned and unsanctioned deceptive communication. *Journal of Nonverbal Behavior*, 22 (3), 189-204

- Frank, M. G. & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research*, 31 (1), 58-75. doi: 10.1080/00909880305377
- Frank, M. G. & Svieteva, E. (2012). Lies worth catching involve both emotion and cognition: Commentary on Vrij and Granhag (2012). *Journal of Applied Research in Memory and Cognition*, 1 (2), 131-133
- Freedman, N. (1972). The analysis of movement behaviour during the clinical interview. In Seigman, A & Pope, B. (Eds), *Studies in Dyadic Communication* (pp.153-187). New York: Pergamon Press
- Fuller, C. M., Biros, D. P., Burgoon, J. & Nunamaker, J. (2013). An examination and Validation of Linguistic Constructs for studying high-stakes deception. *Group Decision and Negotiation*, 22, 117-134. doi:10.1007/s10726-012-9300-z
- Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F. & Scott Winter, A. (2013). Vocal indices of stress: A review. *Journal of Voice*. Advance online publication. doi:10.1016/j.jvoice.2012.12.010
- Gieler, U., Jungblut, F.M., Harth, W., Reimer, C., Niemeier, V., Horing, C.-M. & Kupfer, J. (2008). Shame and touch: Correlations in a study on 140 healthy persons. *Psychotherapeut*, 53 (1), 27-34
- Glass, G.V. & Stanley, J.C. (1970). *Statistical Methods in Education and Psychology*. Englewood Cliffs, N.J.: Prentice-Hall

- Glenberg, A. M., Schroeder, J. L. & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory and Cognition*, 26 (4), 651-658
- The Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology*, 37 (1), 60-74. doi:10.1177/0022022105282295
- Granhag, P. A. & Stromwall, L. A. (2004) (Eds). *The Detection of Deception in Forensic Contexts* (pp.15-40). Cambridge: Cambridge University Press
- Gretscher, H., Haun, D. B. M., Liebel, K. & Kaminski, J. (2012). Orang-utans rely on orientation cues and egocentric rules when judging others' perspectives in a competitive food task. *Animal Behaviour*, 84 (2), 323-331. doi:10.1016/j.anbehav.2012.04.021
- Hall, M. E. (1986). Detecting Deception in the Voice - An Analysis of the Fundamental Frequency, Syllabic Duration and Amplitude of the Human Voice. Unpublished doctoral dissertation, Michigan State University, East Lansing.
- Hancock, J. T., Thom-Santelli, J., & Ritchie, T. (2004). Deception and design: The impact of communication technology on lying behavior. *CHI Letters*, 6 (1), 129–134
- Hare, B., Call, J. & Tomasello, M. (2006). Chimpanzees deceive a human competitor by Hiding. *Cognition*, 101, 495-514. doi:10.1016/j.cognition.2005.01.011

- Harnsberger, J. D., Hollien, H., Martin, C. A. & Hollien, K. A. (2009). Stress and deception in speech: Evaluating layered voice analysis. *Journal of Forensic Sciences*, 54 (3), 642-650
- Harper, R.G., Wiens, A.N. & Matarazzo, J.D. (1978). *Nonverbal Communication*. New York: Wiley
- Harpster, T., Adams, S. H. and Jarvis, J. P. (2009). Analyzing 911 homicide calls for indicators of guilt or innocence). *Homicide Studies*, 13 (1), 69-93
- Hartwig, M., Granhag, P.A. & Stromwall, L.A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, 13, 213-227.
- Hartwig, M., Granhag, P. A., Stromwall, L. A. & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law & Human Behavior*, 30, 603-619. doi:10.1007/s10979-006-9053
- Hartwig, M., Granhag, P.A., Stromwall, L.A. & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, 29, 469-484.
- Hatfield, E., Cacioppo, J.T. & Rapson, R.L. (1992). Primitive emotional contagion. *Review of Personal and Social Psychology*, 14, 151-177
- Henig, R. M. (2006, February 5). Looking for the lie. *The New York Times*. Retrieved from [http://www.nytimes.com/2006/02/05/magazine/05lying.html?pagewanted=all&\\_r=1&](http://www.nytimes.com/2006/02/05/magazine/05lying.html?pagewanted=all&_r=1&)
- Horvath, F., McCloughlan, J., Weatherman, D. & Slowik, S. (2013). The accuracy of auditors' and layered voice analysis operators' judgements of truth and



- deception during police questioning. *Journal of Forensic Sciences*. Advance online publication. doi:10.1111/1556-4029.12066
- Howell, D. C. (2002). *Statistical methods for psychology* (5<sup>th</sup> ed.). Pacific Grove, CA: Duxbury Thomson Learning, Inc.
- Jack, R. E., Caldara, R. & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, 141 (1), 19-25. doi: 10.1037/a0023463
- Johnstone, T. & Scherer, K. R. (2000). Vocal communication of emotion. In Lewis, M. & Haviland, A. (Eds), *The Handbook of Emotion*. New York: Guilford
- Kassin, S.M., & Gudjonsson, G.H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, 5, 35–67. doi: 10.1111/j.1529-1006.2004.00016.x
- Kassin, S. M. & Fong, C. T. (1999). ‘I’m Innocent!’: Effects of Training on Judgments of Truth and Deception in the Interrogation Room. *Law and Human Behavior*, 23 (5), 499-516
- Kebbell, M. R., Wagstaff, G. F. & Covey, J. (1996) Eyewitness confidence and accuracy: Is it dependent on item difficulty? *British Journal of Psychology*, 87, 653-662.
- Keltner, D. & Buswell, B.N. (1996). Evidence for the distinctness of embarrassment, shame and guilt: A study of recalled antecedents and facial expression of emotion. *Cognition and Emotion*, 10 (2), 155-171

- Koper, R. J. & Sahlman, J. M. (1991). The behavioral correlates of real-world deceptive communication. Paper presented at the Annual Meeting of the International Communication Association. Retrieved online 13 March 2009, from [http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content\\_storage\\_01/0000019b/80/24/18/ae.pdf](http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/24/18/ae.pdf)
- Lazarus, R.S. (1991). *Emotion and Adaptation*. New York: Oxford University Press
- Leal, S., Vrij, A., Mann, S. & Fisher, R.P. (2010). Detecting true and false opinions: The devil's advocate approach as a lie detection aid. *Acta Psychologica*, 134, 323-329
- Levine, T. R., Park, H. S. & McCornack, S. A. (2009). Accuracy in detecting truths and lies: Documenting the 'veracity effect'. *Communication Monographs*, 66 (2), 125-144
- Levine, T.R., Anders, L.N., Banas, J., Baum, K.L., Endo, K., Hu, A.D.S. & Wong, N.C.H. (2000). Norms, expectations and deception: A norm violation model of veracity judgements. *Communication Monographs*, 67 (2), 123-137.
- Mann, S. & Vrij, A. (2006). Police officers' judgement of veracity, tenseness, cognitive load and attempted behavioural control in real-life police interviews. *Psychology, crime & Law*, 12 (3), 307-319
- Mann, S., Vrij, A. & Bull, R. (2002). Suspects, lies and videotape. An analysis of authentic, high-stakes liars. *Law and Human Behaviour*, 26, 365-376

- Mann, S., Vrij, A. & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology*, 89, (1), 137-149. doi:10.1037/0021-9010.89.1.137
- Mann, S., Vrij, A. & Bull, R. (2006). Looking through the eyes of an accurate lie detector. *The Journal of Credibility Assessment and Witness Psychology*, 7 (1), 1-16
- Mann, S., Vrij, A., Leal, S., Granhag, P. A., Warmelink, L., & Forrester, D. (2012). Windows to the soul? Deliberate eye contact as a cue to deceit. *Journal of Nonverbal Behavior*, published online first. doi: 10.1007/s10919-012-0132-y
- Masip, J., Sporer, S.L., Garrido, E. & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law*, 11 (1) 99-122
- McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs*, 59, 1-16.
- Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26, 469-480.
- Miller, G. R. & Stiff, J. B. (1993). *Deceptive Communication*. Newbury Park, CA: Sage
- Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M. & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology*, 20, 750-756

- Murray, I. R. & Arnott, J.L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, 93 (2), 1097-1108
- Newman, M. L., Pennebaker, J. W., Berry, D. S. & Richards, J. M. (2003). Lying words: predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29 (5), 665-675
- Nisbett, R. E. & Wilson, T. D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-260.
- O'Sullivan, M. (2007). Unicorns or Tiger Woods: Are lie detection experts myths or rarities? A response to On Lie Detection 'Wizards' by Bond and Uysal. *Law and Human Behavior*, 31 (1), 117-123. doi: 10.1007/s10979-006-9058-4
- O'Sullivan, M. & Ekman, P. (2004). The wizards of deception detection. In Granhag, P. A. & Stromwall, L. A. (Eds), *The Detection of Deception in Forensic Contexts* (pp.269-286). Cambridge: Cambridge University Press
- O'Sullivan, M., Ekman, P. & Friesen, W. V. (1988). The effect of comparisons on detecting deceit. *Journal of Nonverbal Behaviour*, 12, 203–216.
- Partala, T. & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, 59 (1-2), 185-198. doi:10.1016/S1071-5819(03)00017-X
- Porter, S., Campbell, M. C., Stapleton, J. & Birt, A. R. (2002). The influence of judge, target, and stimulus characteristics on the accuracy of detecting deceit. *Canadian Journal of Behavioural Science*, 34 (3), 172-185

- Porter, S., Doucette, N. L., Woodworth, M., Earle, J. & MacNeil, B. (2008). Halfe the world knowes not how the other halfe lies: Investigation of verbal and non-verbal signs of deception exhibited by criminal offenders and non-offenders. *Legal and Criminological Psychology*, 13, 27-38.  
doi:10.1348/135532507X186653
- Porter, S. and ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science*, 19 (5), 508-514
- Porter, S. & ten Brinke, L. (2010). The truth about lies: What works in detecting high-stakes deception? *Legal and Criminological Psychology*, 15, 57-75
- Porter, S. & Yuille, J.C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior*, 20 (4), 443-458
- Reagan-Cirincione, P. (1994). Improving the accuracy of group judgment: A process intervention combining group facilitation, social judgment analysis, and information technology. *Organizational Behavior and Human Decision Processes*, 58, 246–270.
- Reinhard, M. (2010). Need for cognition and the process of lie detection. *Journal of Experimental Social Psychology*, 46 (6), 961-971. doi:  
10.1016/j.jesp.2010.06.002
- Rizzolatti, G., Fogassi, L. & Gallese, V. (2006). Mirrors in the mind. *Scientific American*, 295 (5), 54-61.

- Rycyna, C. C., Champion, C. D., & Kelly, A. E. (2009). First impressions after various types of deception: Less favourable following expectancy violation. *Basic and Applied Psychology*, 31, 40-48. doi: 10.1080/01973530802659851
- Salfatti, C.G. (2003). Offender interaction with victims in homicide: A multidimensional analysis of frequencies in crime scene behaviors. *Journal of Interpersonal Violence*. 18, 499-512.
- Sapir, A. (1987). *Scientific Content Analysis (SCAN)*. Laboratory of Scientific Interrogation. Phoenix, AZ.
- Serota, K. B., Levine, T. R. & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research*, 36, 2-25. doi:10.1111/j.1468-2958.2009.01366.x
- Shaw, J., Porter, S., & ten Brinke, L. (2013). Catching liars: Training mental health and legal professionals to detect high-stakes lies. *Journal of Forensic Psychiatry & Psychology*, published online.  
doi:10.1080/14789949.2012.752025
- Shuy, R.W. (1998). *The Language of Confession, Interrogation and Deception*. Thousand Oaks, California: Sage Publications Inc
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J. & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303 (5661), 1157-1162.
- Sporer, S. L. & Schwandt, B (2006). Paraverbal indicators of deception: A meta-analytic Synthesis. *Applied Cognitive Psychology*, 20, 421-446

- Steller, M. & Kohnken, G. (1989). Criteria-Based Content Analysis. In D.C. Raskin (Ed.), *Psychological Methods in Criminal Investigation and Evidence* (p217-245). New York: Springer
- Stromwall, L.A., Hartwig, M. & Granhag, P.A. (2006). To act truthfully: Nonverbal behavior and strategies during a police interrogation. *Psychology, Crime & Law*, 12, 207-219.
- Stromwall, L.A., Granhag, P.A., Hartwig, M., Wolf, A., Roos af Hjelmsater, E. & Vrij, A. (2006). Will the truth set them free? Suspects' strategies during investigative interviews. Paper presented at the 3rd meeting of the Nordic Network for research on Psychology and Law, Turku, Finland.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using Multivariate Analysis*. Boston: Allyn & Bacon
- ten Brinke, L., & Porter, S. (2012). Cry me a river: Identifying the behavioral consequences of extremely high-stakes interpersonal deception. *Law and Human Behavior* (in press)
- Turner, R.E., Edgley, C, & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology*, 11, 69-89.
- Van Koppen, P. J. (2012). Deception detection in police interrogations: Closing in on the context of criminal investigations. *Journal of Applied Research in Memory and Cognition*, 1 (2), 124-125. doi:10.1016/j.jarmac.2012.04.005

- Villar, G., Arciuli, J., & Mallard, D. (2011). Use of “um” in the deceptive speech of a convicted murderer. *Applied Psycholinguistics*, 33 (1), 83-95. doi: 10.1017/S0142716411000117
- Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *Journal of Social Psychology*, 133 (5), 601-610
- Vrij, A. (2004). Guidelines to catch a liar. In Granhag, P. A. & Stromwall, L. A. (Eds), *The Detection of Deception in Forensic Contexts* (pp.287-314). Cambridge: Cambridge University Press
- Vrij, A. (2005). Criteria base content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy & Law*, 11 (1) 3-41
- Vrij, A. (2008). Detecting lies and deceit: Pitfalls and opportunities. (2nd ed.). Chichester, England: Wiley.
- Vrij, A., Akehurst, L., Soukara, S. & Bull, R. (2004). Detecting deceit via analysis of verbal and nonverbal behaviour in children and adults. *Human Communication Research*, 30 (1), 8-41.
- Vrij, A., Edward, K., Roberts, K. P. & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behaviour. *Journal of Nonverbal Behavior*, 24 (4), 239-263



- Vrij, A. & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, 1 (2), 110-117. doi:10.1016/j.jarmac.2012.02.004
- Vrij, A. & Mann, S. (2001a). Telling and detecting lies in a high-stake situation: the case of a convicted murderer. *Applied Cognitive Psychology*, 15, 187-203
- Vrij, A. & Mann, S. (2001b). Who killed my relative? Police officers' ability to detect real-life, high-stakes lies. *Psychology, Crime and Law*, 7, 119-132
- Vrij, A. & Mann, S. (2004). Detecting deception: The benefit of looking at a combination of behavioural, auditory and speech content related cues in a systematic manner. *Group Decision and Negotiations*, 13, 61-79
- Vrij, A., Mann, S., Fisher, R. P., Leal, S., Milne, R. & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, 32, (3), 253-265. doi:10.1007/s10979-007-9103-y
- Vrij, A., Mann, S., Leal, S. & Fisher, R. (2010). 'Look into my eyes': can an instruction to maintain eye contact facilitate lie detection? *Psychology, Crime & Law*, 16, (4), 327-348. doi:10.1080/10683160902740633
- Vrij, A., Mann, S., Robbins, E. & Robinson, M. (2006). Police officers ability to detect deception in high stakes situations and in repeated lie detection tests. *Applied Cognitive Psychology*, 20, 741-755. doi:10.1002/acp.1200
- Walczyk, J. J., Griffith, D. A., Yates, R., Visconte, S. R., Simoneaux, B. & Harris, R.L. (2012). Lie detection by inducing cognitive load: Eye movements and

- other cues to the false answers of “witnesses” to crimes. *Criminal Justice and Behavior*, 39 (7), 887-909. doi:10.1177/0093854812437014
- Wagstaff, G. F., Wheatcroft, J., Cole, J., Brunas-Wagstaff, J., Blackmore, V., & Pilkington, A. (2008). Some cognitive and neuropsychological aspects of social inhibition and facilitation. *European Journal of Cognitive Psychology*, 20, 828-846. doi: 10.1080/09541440701469749
- Wardle, M. and D. Gloss. (1982). Effects of Lying and Conformity on Decision-Making Behavior. *Psychological Reports*, 51, 871–877.
- Wheatcroft, J. M., Wagstaff, G. F. & Kebbell, M. R. (2004). The influence of courtroom questioning style on actual and perceived eyewitness confidence and accuracy. *Legal & Criminological Psychology*, 9, 1-20.
- Wheeler, B. C. (2009). Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings of the Royal Society: Biological Sciences*, 276, 3013-3018. doi:10.1098/rspb.2009.0544
- Wild, B., Erb, M. & Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Research*, 102, 109-124
- Wright Whelan, C. (2009). The art of deception and the science of detection: Cues to deceit in real life, high-stakes situations. Unpublished Masters dissertation, University of Liverpool
- Zhou, L., Burgoon, J.K., Twitchell, D.P., Qin, T. & Nunamaker Jr., J.F. (2004). A comparison of classification methods for predicting deception in computer-

mediated communication. *Journal of Management Information Systems*,  
20 (4), 139-165

Zuckerman, M., DePaulo, B.M & Rosenthal, R. (1981). Verbal and nonverbal  
communication in deception. In L. Berkowitz (Ed), *Advances in  
Experimental Social Psychology* (Vol. 14, pp.1-59). New York: Academic  
Press

# *Appendix 1*

## **Participant information sheets, studies 2 and 4:**

### **PARTICIPANT INFORMATION SHEET**

**Study title:** High stakes lies: Identifying and using cues to deception in false appeals for missing and murdered relatives

**Researchers:** Prof. Graham Wagstaff  
Clea Whelan

You are being invited to participate in a research study. Before you decide whether to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and feel free to ask us if you would like more information or if there is anything that you do not understand. Please also feel free to discuss this with your friends and relatives if you wish. We would like to stress that you do not have to accept this invitation and should only agree to take part if you want to.

Thank you for reading this.

#### **Purpose of study**

This research is concerned with investigating behaviours associated with deception in real life, high-stakes situations, specifically people making appeals for missing or murdered relatives

#### **Why have I been chosen to take part?**

You are being asked to participate because we are using opportunity sampling for this study/ you are a student at the University of Liverpool.

#### **Do I have to take part?**

Participation is entirely voluntary and you may withdraw at any time during the experimental session without having to give a reason, and without detriment to you. If you withdraw after the study has begun we will destroy any data already collected. Under the Data Protection Act, you have the right to access your data upon request. If you choose to withdraw and you are part of the EPR scheme, you will still be awarded EPR points *pro rata* on the amount of time you have spent being tested, with a minimum of one point.

#### **What will happen if I take part?**

If you decide to take part, you will be asked to attend for a viewing session. You will be briefed on the behaviours you will be rating. You will be asked to sit in front of a computer monitor and view a series of short video clips, showing people appealing for help with missing or murdered relatives. For each clip, you will be asked to rate the people featured on a number of pre-specified behaviours. The

clips have a similar content to what may be viewed in a news programme. After each clip has been played, you will be asked to state if you believe the person is being deceptive or truthful.

The experiment will be run by the student researcher, Clea Whelan, and will take place in a laboratory in the School of Psychology, or equivalent setting.

**Are there any risks in taking part?**

There are no anticipated physical or mental risks associated with this study. However, should you feel any discomfort, you should inform the experimenter and if necessary, the procedure will be terminated.

**Are there any benefits in taking part?**

There are no intended benefits in taking part in the experiment.

**What if I am unhappy or if there is a problem?**

If you are unhappy, or if there is a problem, please feel free to let us know by contacting Clea Whelan (clea.whelan@liverpool.ac.uk) and we will try to help. If you remain unhappy or have a complaint which you feel you cannot come to us with then you should contact the Research Governance Officer on 0151 794 8290 (ethics@liv.ac.uk). When contacting the Research Governance Officer, please provide details of the name or description of the study (so that it can be identified), the researcher(s) involved, and the details of the complaint you wish to make.

**Will my participation be kept confidential?**

Your responses on the response sheet will be stored and later analysed. The data will be anonymised and no personal identifying data will be stored. Only the principal investigator and student researcher will have access to the data, which will be stored on their personal computers.

**Will my taking part be covered by an insurance scheme?**

Participants taking part in a University of Liverpool ethically approved study will have cover.

**What will happen to the results of the study?**

The results of the present study may be published at a future date in an academic journal. You may request that the Principal Investigator contact you at a later date to inform you about accessing published data. No one participating in the study will be identifiable in any resulting publications.

**What will happen if I want to stop taking part?**

You may withdraw at any time without having to give a reason. If you decide to withdraw before completing the experiment, any data collected will not be used and will be electronically deleted.

**Who can I contact if I have further questions?**

Please address any immediate concerns or questions to the student researcher. If, after having completed the experiment, you have any further questions please contact the Principal investigator:

Prof. Wagstaff,  
School of Psychology,  
University of Liverpool,  
Eleanor Rathbone Building,  
Bedford Street South,  
Liverpool  
L69 9ZA

## **Participant information sheets, Study 7:**

### **PARTICIPANT INFORMATION SHEET**

**Study title:** High stakes lies: identifying and using cues to deception in false appeals for missing and murdered relatives

**Researchers:** Prof. Graham Wagstaff  
Clea Whelan

You are being invited to participate in a research study. Before you decide whether to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and feel free to ask us if you would like more information or if there is anything that you do not understand. We would like to stress that you do not have to accept this invitation and should only agree to take part if you want to.

#### **Purpose of study**

This research is concerned with investigating behaviour associated with deception in real life, high-stakes situations.

#### **Why have I been chosen to take part?**

You are being asked to participate because you are a police officer with Merseyside Police/a student at the University of Liverpool

#### **Do I have to take part?**

Participation in the experiment is voluntary, and should you agree to participate you are free to withdraw at any point during the experimental session.

#### **What will happen if I take part?**

If you decide to take part, you will be asked to attend for a viewing session. You will be shown a series of approximately 30 short video clips, showing people talking about missing or murdered relatives. For each clip, you will be asked to decide whether you think the person featured is deceptive or honest. At the end of the viewing, you will be asked to note any behaviours that you use to come to your decision. Finally, you will be given a list of behaviours that may be associated with honesty or deception, and asked to note how often you use them when making your decisions. The experiment will be run by the student researcher, Clea Whelan, and will take place in a laboratory in the School of Psychology.

#### **Are there any risks in taking part?**

There are no anticipated physical or mental risks associated with this study. However, should you feel any discomfort, you should inform the experimenter and if necessary, the procedure will be terminated.

**Are there any benefits in taking part?**

There are no intended benefits in taking part in the experiment.

**What if I am unhappy or if there is a problem?**

If you are unhappy, or if there is a problem, please feel free to let us know by contacting Prof. Wagstaff ([gwf@liverpool.ac.uk](mailto:gwf@liverpool.ac.uk)) and we will try to help. If you remain unhappy or have a complaint which you feel you cannot come to us with then you should contact the Research Governance Officer on 0151 794 8290 ([ethics@liv.ac.uk](mailto:ethics@liv.ac.uk)). When contacting the Research Governance Officer, please provide details of the name or description of the study (so that it can be identified), the researcher(s) involved, and the details of the complaint you wish to make.

**Will my participation be kept confidential?**

Your responses will be stored and later analysed. The data will be anonymised and no personal identifying data will be stored. Only the principal investigator and student researcher will have access to the data, which will be securely stored.

**Will my taking part be covered by an insurance scheme?**

Participants taking part in a University of Liverpool ethically approved study will have cover.

**What will happen to the results of the study?**

The results of the present study may be published at a future date in an academic journal. You may request that the Principal Investigator contact you at a later date to inform you about accessing published data. No one participating in the study will be identifiable in any resulting publications.

**What will happen if I want to stop taking part?**

You may withdraw at any time without having to give a reason. If you decide to withdraw before completing the experiment, any data collected will not be used and will be electronically deleted.

**Who can I contact if I have further questions?**

Please address any immediate concerns or questions to the student researcher. If, after having completed the experiment, you have any further questions please contact the Principal investigator:

Prof. Wagstaff,  
School of Psychology,  
University of Liverpool,  
Eleanor Rathbone Building,  
Bedford Street South,  
Liverpool  
L69 8ZA

0151 794 2949

[gwf@liverpool.ac.uk](mailto:gwf@liverpool.ac.uk)

## *Appendix 2*

### **Participant debrief sheets, studies 2 and 4:**

#### **PARTICIPANT DEBRIEFING SHEET**

**Study title:** High stakes lies: Identifying and using cues to deception in false appeals for missing and murdered relatives

**Researchers:** Prof. Graham Wagstaff  
Clea Whelan

Thank you for participating in this research.

#### **Purpose of study**

This research is concerned with investigating behaviours associated with deception in real life, high-stakes situations, specifically people making appeals for missing or murdered relatives. The hypothesis being investigated in this study is that there are clusters of behaviours that are predictive of deception in the specific context of making an appeal for a missing or murdered relative. People making false appeals are expected to display either fake emotion, or a lack of emotion, and all false appealers are expected to display distancing behaviours. False appealers are expected to elicit a negative personal reaction from viewers, and real appealers are expected to elicit a sympathetic personal reaction from viewers. The data that you have provided will be used to test these predictions.

#### **Who can I contact if I have further questions?**

Please address any immediate concerns or questions to the student researcher. If, after having completed the experiment, you have any further questions please contact the Principal investigator:

Prof. Wagstaff,  
School of Psychology,  
University of Liverpool,  
Eleanor Rathbone Building,  
Bedford Street South,  
Liverpool  
L69 3ZA

0151 794 2949  
[gwf@liverpool.ac.uk](mailto:gwf@liverpool.ac.uk)



## Participant debrief sheets, study 7:

### PARTICIPANT DEBRIEFING SHEET

**Study title:** High stakes lies: Identifying and using cues to deception in false appeals for missing and murdered relatives

**Researchers:** Prof. Graham Wagstaff  
Clea Whelan

Thank you for participating in this research.

#### **Purpose of study**

This research is concerned with investigating behaviours associated with deception in real life, high stakes situations, specifically people making appeals for missing or murdered relatives. It is expected that there will be clusters of behaviours that are predictive of deception in the specific context of making an appeal for a missing or murdered relative. The hypothesis being investigated in this study is that police officers will achieve accuracy rates above chance in detecting deception in this high stake context. The data that you have provided will be used to test this prediction.

#### **Who can I contact if I have further questions?**

Please address any immediate concerns or questions to the student researcher. If, after having completed the experiment, you have any further questions please contact the Principal investigator:

Prof. Wagstaff,  
School of Psychology,  
University of Liverpool,  
Eleanor Rathbone Building,  
Bedford Street South,  
Liverpool  
L69 8ZA

0151 794 2949  
[gwf@liverpool.ac.uk](mailto:gwf@liverpool.ac.uk)

## ***Appendix 3***

### **Case summaries of all appeals used as stimulus materials**

#### **Cases 1 – 9, used as stimulus materials in studies 1, 2 and 3.**

#### **Five honest appealers, and four deceptive appealers.**

##### Case 1: honest appealer, Aneta Kluk

Polish student Angelika Kluk was on a working vacation in Glasgow, when she was reported missing on 24 September 2006. Her sister, Aneta Kluk, made a televised appeal for help three days after Angelika was last seen. Angelika's body was found on 29 September under the floorboards of a church, raped, beaten and stabbed. Peter Tobin, a convicted sex offender who worked at the church, was arrested on 1 October and charged with murder later that week. Tobin denied the charges but extensive DNA evidence (including his semen in her vagina, his DNA on the cloth stuffed in to her mouth, his fingerprints on items found with the body, a pair of his jeans soaked in her blood), implicated him in the murder and he was convicted on 4 May 2007 and sentenced to a minimum of 21 years in prison. He has since been convicted of two other murders of young women.

##### Case 2: honest appealer, Beatrice Jones

Moira Jones' badly beaten and raped body was found in a Glasgow park on 29 May 2008. Five days later her mother, Beatrice Jones, made an appeal for help in catching her killer. A local resident alerted police to Marek Hrcar, who had fled to Slovakia.

CCTV footage showed him leading Moira towards the park where she was found dead. He was found with items belonging to Moira, including a camera and jewellery. He was extradited to the UK on 16 July 2008 and charged with murder. DNA evidence included her blood on the jacket he was wearing when arrested in Slovakia, and his semen on her body. At trial he pleaded not guilty, but overwhelming DNA evidence was presented and he was convicted on 8 April 2009 and sentenced to a minimum of 25 years in prison.

#### Case 3: honest appellant, Sandra Boss

Seven year old Reigh Boss was abducted by her father, Clark Rockefeller, on 27 July 2008 during a supervised visit in Boston. Reigh was grabbed by her father and taken to a waiting car, then driven away as a social worker clung to the door. Reigh's mother, Sandra Boss, made a public appeal for her daughter's safe return shortly afterwards. After a week-long international manhunt, Reigh was returned safely to her mother. Clark Rockefeller was convicted of kidnapping on 9 June 2009 and sentenced to five years in prison.

#### Case 4: honest appellant, Paul Bowman

Sally Anne Bowman was raped and murdered in the driveway of her London home in September 2005. Two weeks later her father, Paul Bowman, appealed for help in catching her killer. There was DNA evidence on Sally Anne's body, but no match on the police database. On 15 June 2006 a swab was taken from Mark Dixie, arrested for a minor unrelated offence, and his DNA was found to match that at the scene of Sally Anne's murder. Dental records also matched him to bite marks on her body. He also left fingerprints at the crime scene. He was charged on 29 June and at his trial on

4 February 2008 he pleaded not guilty, claiming he had come across her dead body and had sex with her. He was found guilty on 22 February and sentenced to a minimum of 34 years.

Case 5: honest appealer, Jim Duell

Nineteen year old Tania Nicol was reported missing on 1st November 2006 from her home in Ipswich. Her body was found in a brook on 8<sup>th</sup> December, six days after another woman's body was found in the same area. By 12<sup>th</sup> December, the bodies of five women had been found, all in similar circumstances. Tania's father, Jim Duell, made an appeal on 15<sup>th</sup> December. On 19<sup>th</sup> December Steven Wright was arrested, and charged with all five murders two days later. At his trial in January 2008, Wright pleaded not guilty to murder. Blood from two of the victims was found on a jacket belonging to Wright, Wright's DNA was found on three of the victims, and fibres from his clothes, his sofa and car were found on all five victims. In particular, a black nylon fibre found in Tania's hair matched the floor carpet in the passenger side of Wright's car, and forensic experts testified that car carpets did not shed fibres easily and suggested the victim's head must have had forceful and sustained contact with the car floor. CCTV footage also captured Tania getting in to Wright's car the night she disappeared. Wright was found guilty of all five murders and sentenced to a whole life term imprisonment in February 2008.

Case 6: deceptive appealer, William Walsh

Walsh's wife Leah's abandoned car was found on Long Island on 27 October 2008, and he made several appeals for his missing wife in the following two days. Her body was found two days after she disappeared, and he was linked to the crime when

it emerged that he had been seen at his wife's abandoned car the morning she was reported missing. William Walsh was arrested on October 29<sup>th</sup>. There were inconsistencies in his story, and items missing from Leah's purse were found in his shoe. He admitted to strangling Leah and dumping her body, and provided police with a seven page confession detailing the events that led up to Leah's death, and his subsequent attempts to provide himself with an alibi. He was charged with second degree murder on 30<sup>th</sup> October 2008 and pleaded guilty to the charge and to tampering with evidence in May 2010. In court he told the judge that he had strangled Leah in a rage after she accused him of infidelity, saying "I put her in a chokehold. I covered up my actions after the fact. I wish I hadn't done any of it. I am truly sorry." He was sentenced to 18 years to life in June 2010.

#### Case 7: deceptive appealer, Diane Downs

On May 19 1983, Diane Downs arrived at an Oregon hospital with her three young children in her car, all of whom had been shot. She claimed that she had been car-jacked and a stranger had shot her and her children. One of her daughters died, her son was left paraplegic and her other daughter was permanently paralysed on one side of her body. Witnesses saw her driving her car very slowly (5-7 mph) towards the hospital. Physical evidence (including blood spatter patterns and shells found in her home that had gone through the same gun casing as the bullets fired in to her children) did not support her story but implicated her in the attack and on 24 February 1984 she was arrested and charged with murder and attempted murder. Her surviving daughter testified against her at the trial, at which she pleaded not guilty. She was found guilty of all charges on 17 June 1984 and sentenced to life in prison plus 50 years.

Case 8: deceptive applier, Matthew Gretz

Gretz's wife, Kira Simonian, was found murdered in Minneapolis on 28 June 2007. She had suffered multiple stab wounds and blunt force trauma. Gretz attended a candlelit vigil for his wife in July and spoke publicly about his loss. Neighbours had reported hearing arguing on the night of the murder, and Gretz shouting 'Do you love me'. Substantial DNA evidence on both the victim and on Matthew Gretz implicated him in the murder and he was arrested and charged in September 2007. Half an hour after murdering Kira, Gretz had flown to New York from Minneapolis (where the couple lived), on a business trip. Police found Kira's blood on his watch, hands, shirt, and on the suitcase that he took with him, and he had fresh cuts and bruises. Gretz's DNA was found under Kira's fingernails. After a judge ruled that the physical evidence was admissible, Matthew Gretz pleaded guilty to second-degree intentional murder on 2 June 2008. He was sentenced to 25 years.

Case 9: deceptive applier, Penny Boudreau

Boudreau reported her daughter, Karissa, missing on 27 January 2008 in Nova Scotia. Two days later she made a televised appeal for help in finding her daughter. Karissa's frozen body was found on a riverbank on 9 February. Inconsistencies in her story placed suspicion on Boudreau, and an undercover operation was launched. On 11 June, Boudreau gave a detailed account and re-enactment to an undercover operator of how she strangled her daughter, in the belief that the operator was part of a crime syndicate and would be able to destroy police evidence against her. The account was secretly recorded. Boudreau was arrested and charged with first-degree murder on 14 June 2008, and made a full confession, admitting that she had strangled

her daughter with a piece of twine in a bid to keep her boyfriend. On 30 January 2009 she pled guilty to second-degree murder and was sentenced life.

**Cases 10 – 42, used as stimulus materials in studies 4, 5, 6 and 7.**

**16 deceptive appealers, and 16 honest appealers.**

Case 10: deceptive applier, Brandon Lockett

On 14<sup>th</sup> January 2010, Lockett reported to police that intruders had broken in to his Virginia home, assaulted him, and kidnapped his two year old stepson, Aveion Lewis, for a \$10,000 ransom. Lockett appealed for help a few days after the alleged kidnapping, but inconsistencies in the abduction story quickly became apparent and within a week he admitted that Aveion was already dead before he had been reported missing. Lockett claimed that Aveion had choked to death while eating a bun and the body had been disposed of by a friend whose name he did not know. Aveion's body was discovered two weeks later in a landfill, and medical examiners later found no pastry lodged in Aveion's throat. They found no food in his body at all, and although the exact cause of death could not be established, his body bore scars, severe burns, a long-untreated broken arm and evidence of weight loss drastic enough to have been fatal. Lockett pleaded no contest and was convicted of child abuse and second degree murder, and in a statement to the court said "My actions aren't excusable and were very wrong. I messed up." He was sentenced to 50 years in January 2012.

Case 11: deceptive applier, Clint Horvatt

Summer Smith was murdered on 12<sup>th</sup> December 2008. Her fiancé, Clint Horvatt, claimed that whilst driving together in Florida, they had pulled over when Summer recognised a parked truck, and a man then robbed them and shot Summer in the head. During the week that followed, Horvatt appealed for help in finding the alleged murderer. The investigation revealed inconsistencies in his story, and on December 24<sup>th</sup> Horvatt and his friend, William Foster were charged with Summer's murder, after the murder weapon, Summer's purse and Horvatt's wallet were found buried near Foster's home and phone records linked the two men. Horvatt had staged the murder to appear as a robbery and paid Foster a few hundred dollars to kill Summer. Both Horvatt and Foster were convicted of first degree murder and sentenced to life in 2010.

Case 12: deceptive applier, Darlie Routier

Routier's two sons, Devon and Damon (six and five years old), were murdered in their home in Texas on 6<sup>th</sup> June 1996. She claimed that whilst her husband slept upstairs, she and the children had fallen asleep together in on the sofa in the den, and she had awoken to the intruder assaulting her, and after a struggle he had left the room. She then noticed that her two sons had been murdered, having apparently heard nothing during their attack; her children had held been held face down on the floor and each suffered several deep knife wounds to their torsos. Routier suffered superficial wounds, described by a medical examiner as hesitation wounds, to her neck and arm. A week after the murders, at the children's funeral, Routier spoke to the press about the murder of her children. Investigators discovered extensive inconsistencies between her report and the physical evidence at the crime scene,



including; a screen had been cut open to appear as if the intruder had entered there, but there was no evidence of a body passing through the netting, the dust on the sill was undisturbed, and traces of the screen were found on a knife inside the house; there were no footprints in the wet mulch below the screen or anywhere outside the house, and no blood trail anywhere outside the house; evidence that attempts had been made to clean blood up in some areas of the house; and blood spatter patterns from both her children found on the back of Routier's nightdress indicated blood cast off from the murder weapon in upswing motions. She was arrested on June 18<sup>th</sup> for the murder of Devon and Damon, and tried for the murder of Damon in January 1997, whose DNA, along with that Routier's, was found on a kitchen knife at the scene (the indictment on Devon's death was held in limbo by the prosecution). Routier pleaded not guilty but was convicted and sentenced to death in February 1997. She is currently on death row.

#### Case 13: deceptive appealer, Derek Fleming

Fleming's 23 year old daughter, Linda, was reported missing from her home in Yorkshire on 22<sup>nd</sup> January 1993, and he made a televised appeal a few days later. Her body was found on 1<sup>st</sup> February, and within a week Fleming was charged with her murder after her blood was found in his home, the boot of his car, and on his shoes. At his trial in January 1994 he admitted hitting Linda eleven times with a lump hammer during an argument, wrapping her body in a polythene bag and then driving her body to moorland and dumping her in a ditch. He denied murder but admitted manslaughter. Fleming was found guilty and sentenced to life. He was released in 2007.

#### Case 14: deceptive appealer, Fadi Nasri

Nasri's wife, Nisha Patel-Nasri, was found in their London front garden bleeding to death from a stab wound, on 11 May 2006. Shortly afterwards, Nasri made a televised appeal for help to catch his wife's murderer. Four months later the murder weapon was found in a drain and CCTV footage identified the car from which the knife was discarded. The driver and another man, paid by Fadi Nasri to kill Nisha, were linked to Fadi Nasri by phone records. On 1 March 2007 he was charged with murder and pleaded not guilty at his trial. He was convicted on 28 May 2008 and sentenced to life imprisonment.

#### Case 15: deceptive appealer, Gordon Wardell

Carol Wardell's body was found on 12<sup>th</sup> September in a lay-by in Warwickshire; she had been asphyxiated. Her husband, Gordon Wardell, was discovered at their home, bound and gagged. He claimed that he returned to his home the previous evening to find that Carol was being held captive by a gang of five intruders, who tied him up, put a cloth over his face to administer a drug which rendered him unconscious for 10 hours, and took his wife to gain access to the building society in which she worked. Four days later, Wardell appeared at a press conference and made an appeal. Wardell was arrested on 20<sup>th</sup> October after inconsistencies emerged in his story, including medical evidence that there was no drug that could be administered in such a way that would render a person unconscious for more than a few minutes. There were no glove or fingerprints from the alleged gang in the Wardell's house, and post-mortem evidence showed that it was impossible for Carol to have been alive at the time that Wardell claimed he had arrived home the night before her body was found. At his

trial, Wardell denied murder, but was convicted and sentenced to 18 years in December 1995.

Case 16: deceptive applier, Graham Alderton

On 22<sup>nd</sup> February 1994, a fire at Alderton's home, in county Durham, killed his three children, Charlotte, Amy & Dean. Fire services determined that the fire had been started deliberately, and Alderton appealed for help in catching the culprit within two weeks of the fire. He was arrested the following day, the 11<sup>th</sup> March 1994, after tests revealed that the fire had been started inside the house. In 1995 he was cleared of murder but convicted of manslaughter and sentenced to 10 years. He has been released.

Case 17: deceptive applier, Karen Matthews

On 19<sup>th</sup> February 2008, Matthews' daughter, Shannon, was reported missing in West Yorkshire. She was found, alive, on 14<sup>th</sup> March a few miles from her home, in the home of Michael Donovan, the uncle of Matthews' boyfriend. Shannon had been drugged and tied up. On 8<sup>th</sup> April 2008, Matthews was charged with child neglect and perverting the course of justice after admitting to a police officer that she had known where Shannon was all along. Donovan admitted that he and Matthews had planned the kidnapping for financial gain, by making money from selling their story to newspapers, and from reward money. A joint trial of Matthews and Donovan in November 2008 concluded with both being found guilty of kidnapping, false imprisonment, and perverting the course of justice. Matthews was sentenced to eight years, and has been released.

Case 18: deceptive applier, Kristi Abrahams

Abrahams' daughter, Keisha Weipart, was reported missing on 1<sup>st</sup> August 2010 in New South Wales. Abrahams made an appeal two days later. It emerged that Keisha had not been seen by anyone outside the immediate family (Abrahams and her partner, Robert Smith) for three weeks before her reported disappearance. Abrahams and Smith were put under police surveillance, and on 22<sup>nd</sup> April 2011, on Keisha's birthday, they travelled to a remote location, at which police found Keisha's buried remains. Abrahams and Smith were charged with murder. In December 2011, Smith pled guilty to manslaughter and being accessory after the fact of murder, and admitted that knowing Abrahams had killed Keisha, he had received, harboured, maintained and assisted her. Abrahams has pleaded guilty to the manslaughter of her daughter, and to interfering with her remains. However, the court has not accepted her plea, and she is to be tried for murder.

Case 19: deceptive applier, Mark Hacking

Hacking reported his wife, Lori, missing on 19<sup>th</sup> July 2004 in Salt Lake City, after she failed to return from an early morning jog. He appealed for help from the public that same day. Lori was 5 weeks pregnant, and the couple were moving to North Carolina where Hacking had claimed he had been accepted to study medicine, having recently graduated from college. It was discovered that Hacking had never completed his college degree, and had not applied to study at the medical school. Lori's blood was found on a knife in their apartment, on the headboard of the bed, on carpet in the apartment, and in Lori's car. On the morning of Lori's disappearance, at the time Hacking claimed he had been out searching for her, a receipt showed that he was actually buying a new mattress, and the couple's mattress was found in a trash bin

near the building where Mark worked. Lori's car keys were found in the couple's apartment, although her car was at the park where she usually drove to jog, and the seat was adjusted for a much taller driver than Lori. On July 25, Hacking's brothers stated to police that he had confessed the murder to them, apparently having shot Lori in the head, wrapped her body in the bloodied part of the mattress and some garbage bags, and disposed of the body in a dumpster. Video footage of Hacking dumping Lori's body and driving her car to the park where she usually jogged was later discovered. On 2<sup>nd</sup> August, Hacking was arrested on suspicion of murdering his wife. Hacking was charged with first degree murder on 9<sup>th</sup> August, and Lori's remains were found on 1<sup>st</sup> October at a county landfill. On 15<sup>th</sup> April 2005, Hacking pleaded guilty to murder, admitting in court that he had shot Lori in the head. In June 2006 Hacking was sentenced to six years to life.

#### Case 20: deceptive applier, Michael Gifford-Hull

Gifford-Hull reported his wife, Kirsi, missing from their home in Winchester on 17th September 2005, and appealed on 5<sup>th</sup> October for her to return home. After her body was found in woodland on 6<sup>th</sup> October, Gifford-Hull admitted killing his wife and fabricating evidence that she had left him, but claimed that she was killed accidentally in their home. In court, he denied murder and claimed that Kirsi's death was an accident that had occurred when she attacked him with her crutches during a row over his use of prostitutes. Kirsi had been strangled, and Gifford-Hull admitted stashing her body under a bed, before throwing it out of a first floor window and hiding it in a garage. Post mortem evidence revealed that Kirsi had multiple fractures to her back and ribs, consistent with being thrown out of a window. He later took Kirsi's body to woodland, to which he returned twice in attempts to bury the body,

something he described to police as 'blooming hard work'. He was convicted of her murder in November 2006 and sentenced to 17 years.

Case 21: deceptive appealer, Paul Brown

On 28<sup>th</sup> May 1984 in London, Brown claimed that his car had been stolen, with his baby daughter, Louise, in the back. Louise was two weeks old and had Down's syndrome. The car was discovered nearby two hours later, with no sign of Louise. Louise's body was never found. On 18<sup>th</sup> June, Brown was charged with murder, after police investigations indicated that Brown had killed Louise in the early hours of 27<sup>th</sup> May after drinking heavily, and his brother had helped him dispose of the body in Brighton that same day. Brown denied murder and his defence focused on the diminished responsibility of a father discovering that his child has a disability. He was convicted of manslaughter and perverting the course of justice, and sentenced to five years in March 1985.

Case 22: deceptive appealer, Paul Dyson

Dyson reported his girlfriend, Joanne Nelson, missing from their home in Hull on 14<sup>th</sup> February 2005. Two days later he appealed for information about her. He was arrested on 18<sup>th</sup> February after Dyson admitted to a friend that he had strangled Joanne in a row about laundry, and the friend told Dyson's mother, who contacted police. CCTV footage, shortly after Joanne's murder, showed Dyson buying gloves and refuse sacks (in which Joanne's body was found), and filling Joanne's car up with petrol (Joanne was in the boot), before driving it to the woodland where he dumped her body. Dyson confessed to killing Joanne and was charged with murder on 21<sup>st</sup> February. A description of the location from Dyson where Joanne's body had

been dumped, and pollen traces on items belonging to Dyson, directed police to the area of woodland in which Joanne's body was found, on 24<sup>th</sup> March. A post mortem confirmed that she had been strangled. Dyson pleaded guilty to murder on 7<sup>th</sup> November 2005, admitting that he had strangled Joanne, and was sentenced to 16 years.

#### Case 23: deceptive applier, Scott Peterson

Peterson's wife, Laci, was reported missing by her parents on 24<sup>th</sup> December 2002 in California. Laci was 8 months pregnant with their son, Conor. Peterson appeared before the press several times in January 2003. Police investigations revealed inconsistencies in Peterson's account, and that he had been having numerous extra-marital affairs. One of his mistresses contacted police on 30<sup>th</sup> December when she discovered that Peterson was married and his wife was missing, as he had told her two weeks before Laci's disappearance that he was a recent widower. She agreed to her phone conversations with Peterson being recorded by police, and later testified against him at trial. On 14<sup>th</sup> April 2003, the body of a foetus was discovered on a shore, and the following day the body of a recently pregnant woman was washed ashore a mile away. The bodies were found 90 miles from the Peterson home, but just five miles from where Peterson claimed that he had been fishing the day of Laci's disappearance. The woman's body was decapitated, the forearms were missing, parts of the legs were missing, there was duct tape on her thighs, and reports from the medical examiner stated that several ribs were cracked at or around the time of death. DNA tests confirmed that the bodies were those of Conor and Laci. Peterson was arrested on 18<sup>th</sup> April 2003, apparently about to flee the country. Strands of Laci's hair were found embedded in a pair of pliers on Peterson's boat.

Peterson was convicted on 12<sup>th</sup> November 2004 of the first degree murder of Laci, and the second degree murder of Conor and was sentenced to death. He is currently on death row.

Case 24: deceptive appealer, Susan Smith

On 25<sup>th</sup> October 1994, Smith reported to police in South Carolina that she had been car-jacked at an intersection, and the perpetrator had driven away with her two young sons, Michael and Alex, in the car. Smith appealed in the following week. Police were suspicious of Smith's account almost immediately, as the red light at the intersection at which she claimed to have stopped and been carjacked, was only triggered if there was a vehicle on the cross street, but Smith claimed that there were no other vehicles present and therefore no witnesses to the crime. Nine days after Michael and Alex had been reported missing, on November 3<sup>rd</sup>, Smith confessed that she had rolled her car in to a lake with her children inside it. Police had already searched the lake Smith told them about, but she was able to tell investigators that the car was much further out in the lake than they had expected, and gave them the exact distance it had travelled before it sank. The car with the bodies of Michael and Alex in it was found where Smith had indicated they would be, and post mortems revealed that the boys had drowned. At trial, Smith admitted killing her children, and on 22<sup>nd</sup> July 1995, she was convicted of murdering Michael & Alex and was sentenced to life.

Case 25: deceptive appealer, Tracie Andrews

On 1<sup>st</sup> December 1996, Andrews claimed that her fiancé, Lee Harvey, had been murdered in a road rage incident in Worcester. Lee had been stabbed and had his



throat cut. Andrews appealed on 3<sup>rd</sup> December. Inconsistencies in her account (for example, the road on which she claimed that the other car involved in the alleged road rage incident had overtaken Lee's car, was single track), and witness accounts that contradicted Andrews's story, led to Andrews' arrest on 7<sup>th</sup> December. Forensic evidence against Andrews included a bloodstain on the inside of her ankle boot that was consistent with the knife that was used to kill Lee, blood spatters on her clothing that indicated she had been very close to him as he was stabbed, strands of her hair clutched in Lee's hand, and a further clump of hair nearby indicating it had been pulled out in a struggle. Andrews pleaded not guilty at trial, but was convicted of murder on 29<sup>th</sup> July 1997 and sentenced to life. She appealed the sentence, but in April 1999 admitted killing Lee in a letter to her solicitor and to a national newspaper, claiming that it was self-defence and that she had lost control. She was released in 2011.

#### Case 26: honest applier, Alan Symes

On 5<sup>th</sup> October 2009, the mother of two year old Aisling Symes reported her missing from their home in New Zealand. Aisling had been at home with her mother and her five year old sister, Caitlin. The last known person to see Aisling was Caitlin, as they had been playing together in the garden of their home. It was initially thought that Aisling had been abducted, and Aisling's father, Alan, made an appeal on 8<sup>th</sup> October. On 12<sup>th</sup> October a body was found 36 metres down a stormwater drain near the Symes' house, which was confirmed to be Aisling the following day. A post mortem examination determined that Aisling had drowned, and found no sign of any other injury. The manhole cover through which Aisling had fallen was found to be faulty (susceptible to dislodging in heavy rain) and without a safety grate, and had

been due to be replaced several years previously. The city council had received numerous complaints about the safety of the manhole cover from residents, who stated that it lifted after heavy rain. At the inquest into Aisling's death, evidence showed that water pressure had dislodged the manhole cover (there had been heavy rain), and that the cover flipped over when stood on. Aisling's death was ruled an accident, and the coroner made numerous recommendations regarding the safety of manhole covers.

#### Case 27: honest applier, Carrie McGonigle

14 year old Amber DuBois disappeared on 13<sup>th</sup> February 2009 close to her High School in California. Her mother, Carrie McGonigle, appealed in the following week. Amber's skeletal remains were found a year later after John Gardner, who had been arrested in connection with the murder of another local teenage girl, led police to her grave. On 16<sup>th</sup> April 2010, Gardner admitted raping, stabbing and killing Amber, and pleaded guilty to the murder and rape of Amber and Chelsea King. He was sentenced to life without the possibility of parole. In a subsequent TV programme about the murders, Gardner said 'I never want to be let out. I will kill. I know I will.'

#### Case 28: honest applier, Brent King

17 year old Chelsea King disappeared in California on 25<sup>th</sup> February 2010 whilst jogging. Her father, Brent King, made an appeal in the days following her disappearance. On 28<sup>th</sup> February John Gardner was arrested after his DNA was found to match a sample taken from discarded clothing belonging to Chelsea. Her body was discovered on 2<sup>nd</sup> March buried near to where her clothing had been found. Gardner

admitted raping and strangling Chelsea. On 16<sup>th</sup> April 2010, Gardner pleaded guilty to the murder and rape of Chelsea and another local teenager, Amber Dubois. He was sentenced to life without the possibility of parole. In a subsequent TV programme about the murders, Gardner said 'I never want to be let out. I will kill. I know I will.'

Case 29: honest appealer, Gillian McKenna

Graham McKenna and his 14 year old son, Michael, were both stabbed in an attack in Merseyside on 10 January 2009. Graham died of his injuries. His wife, Gillian McKenna, made an appeal on 12<sup>th</sup> February. On 25<sup>th</sup> February, Gary Finlay, who had been the subject of a police manhunt since the attack, handed himself in to police. Finlay was the ex-husband of Gillian's sister, and numerous witnesses testified to his repeated threats of killing Graham and Gillian, who were due to testify against him in a custody hearing. Numerous witnesses saw Finlay stab Graham 18 times, and then stab Michael 12 times, stopping only when members of the public intervened. Finlay then returned to his car and telephoned his sister minutes later, confessing to the attack. Finlay claimed that the attack had not been premeditated, that it was coincidence that he had come across the McKennas, that he had acted in an explosion of rage, and that the knife he used in the attack was in his car because of a previous fishing trip. Finlay pled guilty to murder and attempted murder, and was convicted on 10<sup>th</sup> July 2009 and sentenced to two life sentences.

Case 30: honest appealer, Nick Cook

20 year old Joseph Cook was murdered in Leeds on 30<sup>th</sup> August 2009. His father, Nick Cook, appealed on 10<sup>th</sup> September. Joseph had been stabbed 15 times in an apparent burglary at his home. Gareth Brear was arrested on 23<sup>rd</sup> September when he

aroused the suspicion of staff at a homeless project by repeatedly mentioning the murder. He confessed to police, claiming that he had been high on drugs and alcohol and had burgled the house to fund his drug addiction .At trial, Brear admitted to breaking in to the house through a basement window, stabbing Joseph, ransacking the house, setting light to a mattress in an attempt to destroy evidence, and stealing a BMX bike, which he went on to sell for £10, although he denied bringing the murder weapon with him. On 8<sup>th</sup> March 2010, Brear pled guilty to murder and was sentenced to 26 years.

Case 31: honest appealer, Scott Geschke

Keri Geschke disappeared on 8<sup>th</sup> June 2010 from her home in Wisconsin, and her husband, Scott, appealed in the following days. Keri was found alive and well four days later, when she telephoned Scott. She had gone to the family's hunting trailer for some time alone, but had become lost and then her vehicle had broken down. She left and returned voluntarily.

Case 32: honest appealer, Bernie Sousa

On 21<sup>st</sup> May 2010, Kyleigh Sousa died in Arizona when she was dragged by a car in an attempted robbery of her bag in a car park. Her hands were entangled in her bag straps, and she was dragged 30 feet, suffering a fractured skull and detached artery. A friend who was with her when the event occurred, Franco Hernandez, testified that the car had pulled up alongside them, they had chatted with the occupants, and then the driver leapt out and grabbed Kyleigh's bag. Kyleigh's brother, Bernie Sousa, appealed for help in finding the perpetrator within the following month. Joseluis Marquez had been identified driving the rental vehicle involved in Kyleigh's death in

an earlier speed camera photo, and initially offered an alibi that was later disproven. Police were then contacted by one of the other occupants of the car, who stated that Marquez had been the driver of the car that night. Marquez was arrested when police were able to track down the rental car, and charged with first degree murder December 2010. Police were able to identify the four other occupants of the car, who all named Marquez as the driver and the person who grabbed Kyleigh's bag. Marquez was convicted of first degree murder on 22<sup>nd</sup> October 2012.

#### Case 33: honest applier, Tony Ritch

Lavern Ritch was killed on 12<sup>th</sup> August 2007 in New Jersey, as he tried to intervene in an altercation between two men who were strangers to him. Lavern was stabbed once through the heart by Robert Davies, according to numerous witnesses, who fled the scene but was arrested 11 days later. Lavern's brother and cousin flew from Wales to New Jersey within days of Lavern's death and made an appeal. At trial in January 2011, Davies admitted killing Lavern, claiming that he acted in self-defence as he thought that Lavern was a gang member chasing him down the street. Davies was convicted of reckless manslaughter and sentenced to 21 years. At the sentencing, Davies apologised to Ritch's family.

#### Case 34: honest applier, Kathleen Kirby-Jones

18 year old Michael Jones was tortured and murdered in London on 3<sup>rd</sup> March 2008. His mother, Kathleen Kirby-Jones, returned home from work to find Michael bound in parcel tape, bludgeoned, stabbed and mutilated. Kathleen heard that somebody was still in the house, and barricaded herself in to the room with Michael as the perpetrator tried to force his way in to the room. Kathleen called police, who had to

break in to the house as Kathleen was too afraid to leave the room. The perpetrator, who had been waiting in the house for several hours for Kathleen to return, and had removed light bulbs so that Kathleen would not be able to see, fled before police arrived. Michael did not survive the attack, and Kathleen appealed the following week. Post mortem evidence indicated that the attack had occurred while Kathleen was at work. A car linked to the attack by CCTV footage was traced as belonging to Gerard Paul, with whom Kathleen had been in a relationship, which she had ended the month before. Paul was arrested and charged on 20<sup>th</sup> March and at his trial in December in 2008 evidence was presented that Paul believed that Michael stood in the way of his relationship with Kathleen. Paul was convicted of murder in December 2008 and sentenced to life.

#### Case 35: honest applier, Garry Rentz

According to Nancy Cooper's husband, Brad Cooper, Nancy went for a jog at 7am on 12<sup>th</sup> July 2008 in North Carolina and did not return. A friend reported Nancy missing at noon that day, and her father, Garry Rentz, flew in from Canada and made an appeal the following day. Nancy's body was discovered on 14<sup>th</sup> July, 3 miles from her home. A medical examiner determined that she had been strangled. Initial statements made by Cooper to investigators were inconsistent with the evidence, and with statements that he made later. On 27<sup>th</sup> October 2008, Cooper was arrested and charged with Nancy's murder. Analysis of the search history on Cooper's computer from the days before his wife disappeared revealed multiple zoomed-in satellite images of the exact, remote area where his wife's body was found, copies of emails indicating that Nancy was leaving him, and evidence of an affair that he was having. Soil samples from Cooper's shoes matched soil at the location of Nancy's body, and

tire tracks and footprints further linked Cooper to the crime scene. Cooper was convicted of first degree murder in May 2011 and sentenced to life.

Case 36: honest applier, Ann Szary

85 year old Patricia Thompson's body was found on 5th September 2009 in her home in Teesside. She had been robbed, strangled, and her home set on fire. Patricia's telephone had been used to call a taxi shortly before firefighters arrived, and a man carrying two suitcases had been taken to a local address. Her daughter, Ann Szary appealed the following week. Steven Hodgson was charged with Patricia's murder on 17<sup>th</sup> September. Hodgson's DNA was found under Patricia's fingernails, satellite images and maps of the area where Patricia lived had been accessed on Hodgson's computer two days before the murder, CCTV showed him using her bank card after her death, and he pawned her belongings. Hodgson was convicted of murder and sentenced to life in May 2010.

Case 37: honest applier, Robert Edwards

15 year old Rosemary Edwards disappeared from her home in Hampshire on 5<sup>th</sup> September 2007, after a family argument following her parents' discovery that she had lied about being sacked from a job. A forensic examination of the Edwards' house revealed no evidence of foul play. Her father, Robert Edwards, made an appeal a few days after Rosemary's disappearance. Rosemary's body was found 18 days after her disappearance, hanging from a tree in remote woodland 10 miles from her home. The internet trail on Rosemary's computer, and text messages from her phone, revealed that she had been suffering from depression. Messages sent to four friends the night before she disappeared expressed Rosemary's intention to commit suicide

somewhere that she would not be found. Police reported that the death was not suspicious, and that they were not looking for anybody else in connection to the death. At the inquest in to Rosemary's death in November 2007, the coroner recorded a verdict of suicide.

Case 38: honest applier, Maria Chavez

Eight year old Sandra Cantu was reported missing from her home in California by her mother, Maria Chavez, on 27<sup>th</sup> March 2009. Chavez made an appeal in the week that followed. A neighbour, Melissa Huckaby, reported to police that a suitcase had been stolen from her, and that she had found a note indicating where Sandra's body was, and that it was in a stolen suitcase. Sandra's body was found in Huckaby's suitcase in a pond, on 6<sup>th</sup> April. On 10<sup>th</sup> April, Huckaby was arrested and charged with the kidnap, rape, and murder of Sandra. Sandra had been force-fed prescription drugs and rubbing alcohol, sexually assaulted and battered with a rolling pin, and asphyxiated with a cloth knotted in to a noose. CCTV footage showed Sandra in Huckaby's car at the time she disappeared, leaving the trailer park where they lived and turning towards the nearby church at which Huckaby's grandfather was a minister. The rolling pin used to assault Sandra was found in the church with Sandra's blood on it, as were blood spatters and a half-empty bottle of rubbing alcohol. CCTV footage from an hour later showed Huckaby driving away from the church, and eyewitnesses saw Huckaby at the pond where Sandra's body was found shortly after this time. A notebook was found in Huckaby's trailer with indentations that matched the note she claimed to have found, indicating the location of the body and that it was in a suitcase. Huckaby confessed to police that she had been involved in Sandra's death, but initially claimed that it was an accident, that Sandra had



climbed in to the suitcase during a game of hide and seek. At trial, Huckaby pleaded guilty to first degree murder and kidnapping of Sandra, and was sentenced to life without the possibility of parole in June 2010.

Case 39: honest appealer, Sara Payne

Eight year old Sarah Payne was reported missing on 1st July 2000 from near her grandparents' home in West Sussex. She had been playing in a cornfield with her siblings, who reported seeing a white van when she disappeared. Her mother, Sara Payne, appealed in the week following Sarah's disappearance. Sarah's naked body was found 15 miles away on 17<sup>th</sup> July. On 6<sup>th</sup> February 2001, a local known sex offender, Roy Whiting, was charged with Sarah's abduction and murder, following forensic tests on his white van. DNA tests on a hair found on a red sweatshirt in Whiting's van revealed a one in a billion chance that the hair belonged to someone other than Sarah. DNA tests confirmed that Whiting had worn the sweatshirt. The only item of Sarah's clothing to have been found was one of Sarah's shoes, on the Velcro of which forensic tests revealed there were fibres which matched the same red sweatshirt, and a fibre that matched a cloth found in the van. A ball of hair, soil and vegetation found beside Sarah's body contained fibres that matched the red sweatshirt and socks found in the van, and in a further hair sample, fibres that matched the sweatshirt, socks, and material from the van's passenger seat were found. In the body bag used to remove Sarah's corpse, fibres were found matching the socks and the material of the driver's seat of the van. On 12<sup>th</sup> December 2001, Whiting was convicted of Sarah's abduction and murder and sentenced to life.

Case 40: honest appealer, Bradley Lockhart

Five year old Shaniya Davis was reported missing by her mother, Antoinette Davis, on 10<sup>th</sup> November 2009, from her home in North Carolina. Shaniya's father, Bradley Lockhart, who was separated from Davis and lived in a different state, appealed for his missing daughter within two days of Shaniya's disappearance. CCTV footage from a hotel 30 miles from where she disappeared, clearly showed Shaniya being carried down a hallway by a man, Mario McNeill, the day she was reported missing. When the CCTV footage was made public, McNeill surrendered to police and was charged with kidnapping on 13<sup>th</sup> November. On 14<sup>th</sup> November, Shaniya's mother was charged with human trafficking, child abuse involving prostitution, and filing a false police report. In an affidavit for a warrant to search McNeill's car, investigators stated that McNeill told them he picked Shaniya up in front of her home and drove her to the hotel. Evidence from the hotel confirms that he checked in at 6.11am, and left at 7.33am. Shaniya's body was found in a ditch on 16<sup>th</sup> November, following information provided by McNeill's lawyer. Shaniya had been sexually assaulted and asphyxiated at a time close to when she was reported missing. McNeill was charged with first degree murder and rape of a child, Davis was also charged with murder. Based on a confession made to police by Davis, prosecutors allege that Davis owed McNeill a \$200 drug debt, and offered her daughter to him in lieu of payment. The trial of Mario McNeil is on-going at the time of writing, and his defence do not dispute that he took Shaniya. Bradley Lockhart has never been named a person of interest or a suspect in the case.

#### Case 41: honest applier, Kevin Reape

Sian O'Callaghan was reported missing on 19<sup>th</sup> March 2011, after she failed to return from a night out in Swindon. Her boyfriend, Kevin Reape, appealed in the following

days. CCTV footage showed Sian getting in to Christopher Halliwell's cab (identified through the plate number) at 3am the morning she disappeared. Halliwell was placed under surveillance, and arrested on 24<sup>th</sup> March. He confessed to killing Sian by stabbing her in the head, and also confessed to another murder, that of Becky Godden-Edwards, who had been missing from the area since 2004. He then led police to Sian's body, and to the body of Becky, and was charged with the two murders. Due to a police error, Halliwell's confession to the two murders, and his knowledge of the location of the bodies, were ruled to be not admissible as evidence, and the charge relating to Becky was dropped. A post mortem established that Sian had died from the combined effects of two stab wounds to the head and neck, as well as compression to the neck. Forensic examinations of Halliwell's car found Sian's blood in the boot, and at his trial, Halliwell pleaded guilty to Sian's murder. He admitted to beating Sian, stabbing her twice in the head with a six inch knife, and dumping her body in Savernake Forest. He also admitted to returning to where he had left Sian's body four times, and to moving her remains to a different site. At trial, Halliwell's defence lawyer claimed that Halliwell had stabbed and beaten Sian because he lost his temper when she protested about being taken the wrong way in his cab. He was sentenced to life on 19<sup>th</sup> October 2012.

**‘Filler’ appeals used as stimulus materials in study 7.**

**Two honest appealers and two deceptive appealers.**

Case 42: honest appealer, Gil Harrington

20 year old Morgan Harrington disappeared on 17<sup>th</sup> October 2009 from a concert in Virginia. Her purse and phone were found the following morning in a nearby field. Morgan’s mother, Gil Harrington, appealed in the following two weeks. Morgan’s body was discovered in January 2010 in a hayfield on a remote farm, 10 miles from where she was last seen. DNA evidence from Morgan’s body linked her murder to a suspect in a previous sexual assault in Virginia, in which the attacker was scared away by a passerby. The victim in this case was able to get a good look at the man, enabling police to produce a sketch of the attacker.

Cases 5, 6 and 9, as detailed above

## *Appendix 4*

### **Study 2: Observer Response Sheet**

For all statements use the 5-point rating scale and circle as appropriate

#### **Fake emotion**

1. This person's emotions are fake

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

2. This person is genuinely sad

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

3. This person's facial expressions are fake

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

4. This person's facial expressions are genuine

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

5. This person is pretending to cry

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

6. This person's eyes are genuinely sad

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

7. This person is showing no emotional variation

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

8. This person is putting on a performance

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

9. This person's voice is high-pitched

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

### Lack of emotion

10. This person is showing very little emotion

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

11. This person is containing their emotions

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

12. This person is calm and calculated

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

13. This person is trying to stay calm

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

14. This person has very little facial expression

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

15. This person has used brutal language or detail

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

16. This person has avoided using brutal language or detail

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

17. This person smiles appropriately

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

18. This person is trying to be sensible/rational

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

19. This person's voice is emotionless

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

20. This person is remembering painful events without showing pain

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

21. This person is cold

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

## Distancing

22. This person's appeal is impersonal

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

know/unsure

23. This person's appeal is genuine and heartfelt

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

24. This person is focused on the missing/murdered person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

25. This person is dwelling on other people (not the missing/murdered person)

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

26. This person is showing desperation/urgency

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

27. This person is making general/vague statements about the missing/murdered person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

28. This person is describing the missing/murdered person in positive terms

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

29. This person is listing things/people

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

30. This person is **not** using a prepared speech

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't	Agree	Strongly agree



know/unsure

31. This person is looking away a lot

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

32. This person has credible/believable eye contact

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

33. This person is showing genuine/natural behaviours

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

34. This person is **avoiding** referring to her/himself in the first person (I, me, my etc)

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

35. This person is using modifiers (sort of, kind of, quite, like, maybe etc)

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

### Implausible

36. This person does not seem to have any hope of finding the missing person (if body has already been found, please cross the following box: N/A ☐ )

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

37. This person is hopeful of finding the missing person (if body has already

☐

been found, please cross the following box: N/A )

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

38. What this person is saying does not make sense

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

39. This person is making assumptions about what has happened to the missing/murdered person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

40. It sounds like this person had a dispute with the missing/murdered person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

41. This person is fearful/nervous

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

42. This person seems defensive

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

### Personal reaction

43. I don't feel sympathy for this person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

44. I feel sorry for this person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

45. This person is behaving unnaturally

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

46. This person is behaving as I would expect

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

47. This person is creepy

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

48. This person seems normal

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

49. I dislike this person

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

50. I feel this person's pain

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

51. The person making this appeal is being honest

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Don't know/unsure	Agree	Strongly agree

## *Appendix 5*

### **Study 3: Transcripts of appeals used in Study 3**

#### **Honest appeals**

##### Case 1

If there is anything / a small thing something you think it's not important at all please  
/ get in touch with police anyway. / Anything could help at this very point. We do not  
know.

---

I I just / don't know I I cannot add it up together. Um I'm just very upset about it and  
/ I really hope and and and I pray that Angelika will be back soon and she will be  
back in one piece and and / safe and sound.

##### Case 2

Moira our precious daughter / Grant's big sister / could not have been more loved /  
more widely loved. / Moira couldn't have been more loving / more giving more  
thoughtful. / And our lovely lovely girl has been snatched away from us / in the most  
vicious / and senseless way. / Nothing will ever be the same again. / Nothing can  
change that. / There is nothing that any of you can do / to help us feel better now.

---

Please / think deep and hard. / If you know anything / if you saw anything / heard anything / if you are in the least bit suspicious the least bit suspicious of anyone / please tell the police. / You must tell the police. / Don't wait until the heart is ripped out of another family.

### Case 3

Clark / although things have changed / you will always be Reigh's father and I will always be Reigh's mother. / We both love her dearly / and have only her best interests and well-being at heart. / I ask you now / please please bring Snooks back. / There has to be a better way for us to solve our differences than this way. / I also wanna thank everyone for your help. / And Reigh honey / I love you / I miss you so much and remember you're always a princess.

### Case 4

She had this / wonderful / refreshing air of innocence / although she was also fairly streetwise and sensible. / Um / you know and having only just had her eighteenth birthday / albeit a cliché she / genuinely did have the rest of her life ahead of her.

---

I understand that / it could be someone's brother or son or / nephew or next door neighbour / and I know that obviously there's there's some kind of loyalty there. / But we really do need to catch this man before / before he strikes again.

### Case 5

Tania was a lovely daughter. / Ah she was a caring loving sensitive girl / who would never hurt anyone. / Unfortunately drugs / took her away / in to her own secret world. / A world that neither of us were aware of. / Tania has been taken by someone / who needs to be found. / We ask for anyone who knows this person / or persons / to come forward / and contact the police. / They can't take away our memories / they can't take away our love our fortitude our courage. / Grieve for our daughters but not unnecessarily. / Live your lives through / our departed daughters / as they would want to see us getting on with our lives / and not going around with our heads bowed down.

### **Deceptive appeals**

### Case 6

I want my wife back she has can have / my car she can have everything I just want my wife back. Please just get the picture out that's all we want / alright.

(Reporter: Any idea....)

I got no clue

(Reporter: Did you get a cellphone call? Did you get anything?)

I got a text from her in the morning but that was it.

---

I want my wife back. I miss her more than anything. / Just / need everybody's help. / Everybody / anybody that could have seen anything please just / talk.

(Reporter: Do you have more pictures?)

I'm going all over. I'm gonna try and make more posters. My brothers and sister are out putting up / got a whole bunch a' people out there. / Need somebody if they saw something just / let us know.

### Case 7

There are a lot of memories that / um / I dunno / eh a lot of people when something traumatic happens to them they suppress it immediately. / I / kept those memories because I knew that I was the only person that was gonna be able to tell them what happened when we got to the hospital. / And when I got there the first thing I said was call the doctors second thing was the blood-type third thing was call the cops / because they've gotta / they've gotta find him. And / so I had to remember as much as I could remember.

---

When this man shot my daughter my first reaction was to snap back to my childhood / to the pain that had happened to me / back then / my marriage / my entrapment by society. This man was bigger than me he was stronger than me he had more power because he had a gun. He was in control and I was not. And I had there was nothing I could do / and I stood there and I looked at Christy retching / and / the blood that just kept gushing out of her mouth  
and and / I eh what do you do. / You just stand there trapped. And then / and then the gun kept firing and firing and firing and it it it made / it was monotonous. It just kept going. It was like a slow motion picture.

### Case 8

My wife was murdered / and / every day it gets a little bit harder in some ways but  
it's / kinda days like today / or moments like this / seeing your faces and / seeing  
your support / and knowing that we're all in this together / that makes it just a little  
bit easier.

---

I know that we all appreciated Kira in different ways. / Um / it's not just my loss but  
it's / it's friends' losses / it's students' / losses / it's neighbours' losses. / Going  
forward let's let's tell good stories / about Kira 'cause that's what she would have  
wanted.

### Case 9

I'm just here to reach out to my daughter um / Karissa I just wanna tell you / that you  
have lots of people who love you / and want you home. / Your aunt April is here /  
your mom is here / your dad Shane Verne / your uncle Joey your aunt Crissy / your  
friend Sarah's worried sick / everybody in school your grandmothers / everybody. /  
Please just reach out for someone. / At least call us and let you know you're okay. /  
We all love you. / And if there's anybody out there that knows / have seen her or  
anything please call. / Um / the other thing I wanna say is I wanna thank everyone in  
the community that's been a support / all the businesses we both work for and / the  
community in general. / Um / it's been very comforting. / Nothing can be done to  
make things better but it's comforting to have support. / The main thing is I just want  
somebody to come forward / if not Karissa herself somebody / let me know. / It's  
hard to not know where your kid is.



## *Appendix 6*

**Study 3:** instructions to coders for frequency counts of verbal behaviours in appeal transcripts

### **Indicators of deception**

- Uses brutal language or detail directly about their loved one (eg. ‘murdered’)
- If the victim is missing and a body has not been found: expressions of lack of hope (count each phrase/clause so that emphasis is included)
- Avoids using first person singular pronoun (eg. pronouns disappearing altogether, replacing/switching to first person plural or second/third person pronouns)
- Refers to other people and groups of people (do not include the police or the perpetrator)
- List (3 items, eg. people, actions, emotions). If list is longer than 3 items, count each 3 items as 1 list. Do not include positive descriptions of the victim.
- Modifiers, minimisers and hedging words (eg. Just, kind of, sort of, like, guess, maybe, a little, probably, possibly, quite, some, a bit, dunno)
- Filled pauses (huh, um, ah, er)
- Total number of unfilled pauses
- Pauses in the middle of a clause.
- Speech errors (eg. Incomplete sentences, grammatical errors, sentences that don’t make sense, single word consecutive repetition such as ‘I I’ or ‘and and’. Do not include repetitions that make sense and are used for emphasis,

such as 'please please' or 'lovely lovely'). Do not include normal contractions eg. wanna, gonna

- Repetition of a phrase (2 or more words). Count once for each repetition.
- Sentence or statement that doesn't make sense/is illogical/is unclear
- Sentence or statement that is generic, or about something not relevant to the case

### **Indicators of honesty**

- Avoids using brutal language (eg. saying 'gone' instead of saying 'murdered') about the victim
- If the victim is missing and a body has not been found: expressions of hope (count each phrase/clause so that emphasis is included)
- How many times victim is referred to by name
- How many times victim is referred to in second person (you)
- How many times victim is referred to in third person (he/she/her/they)
- How many times victim is referred to in terms of relationship to appellant (eg. My wife, our daughter, our girl)
- Number of times first person singular is used (I, me, my)
- Number of times first person plural is used (we, us, our)
- Pleas for help (count each phrase/clause so that emphasis is included)
- Words describing victim positively, including terms of endearment and expressions of love for victim. Count each word/phrase.

## Appendix 7

### Study 4: Observer response sheet

This clip features (background of clip inserted here).

Please check this box and do not complete the response sheet if you are familiar with the person featured in the box, or the outcome of this case

☐

Please think about the person in the clip that you have just watched. For the following statements, use the 5-point rating scale and circle your agreement with each statement as appropriate.

#### The following statements are about the emotions of the person in the clip:

**This person's emotions are fake**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person's facial expressions look fake**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is genuinely sad**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person's eyes look genuinely sad**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is cold**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person's statement is genuine and heartfelt**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**There is an urgency in this person's statement**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person seems defensive**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is showing no emotional variation**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is trying to stay calm**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is trying to contain their emotions**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**The following statements are about the behaviour of the person in the clip:**

**This person is putting on a performance**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is pretending to cry**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person seems plausible**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is behaving unnaturally**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person seems like a normal person**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is looking away a lot**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person has credible eye contact (for example, if they are looking down a lot, it is because they are reading a statement)**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**The following statements are about the voice and speech of the person in the clip:**

**This person is hesitant when talking**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is direct and certain in what they are saying**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person is speaking in a personal way, and sounds expressive and involved in what they are saying**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person's voice is quivering with genuine emotion**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**This person's voice is high pitched**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**The following statements are about your personal reaction to the person in the clip:**

**This person seems creepy**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**I feel this person's pain**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**I do NOT feel sympathy for this person**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**I dislike this person**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**I feel sorry for this person**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**The following statements are about the honesty of the person in the clip:**

**The person is being honest**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Strongly disagree	Disagree	Neither agree nor disagree / unsure	Agree	Strongly agree

**If you had to guess whether or not the person in the clip is lying, what would you decide?**

Lying

Truthful

## *Appendix 8*

### **Study 5: Transcripts of appeals**

#### **Deceptive appeals**

##### Case 10

Please bring my son home please / and if anybody out there you know some / call the police and tell them.

Just know / I heard a knock at the door / I went to the back door / and all I remember was boom that's all I remember / when I came / when I came to my wife was standing over 'im /and my son was sh (*unfinished word*) know what I'm saying he was crying / and my daughter she was gagged / upstairs her room / and / that's all I remember /

*Interviewer: and what what about the people*

/ I mean I seen some guys earlier that morning but far as / right then when it happened I don't I don't recall no faces I just / remember blacking / seeing just black / that's all I remember but I remember some guys I had seen earlier that morning / ah / on the side a my house / and also I seen some guys over here in the um food line parking lot / they were standing right out here in food line's parking lot / and I remember those guys but other than that I don't / recall anything else / I just want everything to be okay I hope he alright / I just that's all I hope / because he can't hold his / his own body temperature / he has a lot of ah medical problems

*Interviewer: like what*

He has malrotation / um / meaning ah / he can eat / but and ah eat and eat / but his body won't / ah / absorb the fats that he needs so he don't has the enzymes that he needs so / i mean he he has medical problems/ I mean if you got 'im / make sure you keep 'im warm

#### Case 11

I seen you / d (*unfinished word*) directly / you seen me directly / and um / I know what you look like I'm not gonna / I'm not gonna stop / I'm gonna continue on with the investigators / um / I'm gonna do what I have to do I don't care if I have to talk to es (*unfinished word*) someone every single day every single night it doesn't matter / but um / I know what you look like

#### Case 12

If you knew Devon and Damon / you would know that they're up in heaven / and they're up there having the biggest birthday party / that we could ever imagine / and they wouldn't want us / to be down here being sad even though our hearts are breaking

I didn't feel anything because I was in shock / but I wasn't thinking about me / all I was thinking about was trying to save the babies I mean Darin and I tried to save the babies but / it was too late the babies were gone / but we tried / we tried and we have to live with that forever



### Case 13

Linda / wherever you are / or for whatever reason / please come home / or just ring  
/ we all love you / and we want you home with us

### Case 14

Mm she had a good heart / always very very bubbly / always willing to help  
everyone / hard working

Everyone's grieving and ah missing her very much / still in shock

Obviously someone's got a guilty conscience um / they'll be worrying about what  
they done or shocked or maybe it was an accident or mistake or or or whatever /  
you know ah / but someone has gotta know something / who lives around them / a  
neighbour / or seen some blood or / someone acting suspiciously or nervous / just  
might not be important but / just give us a call and let us know

### Case 15

Then as I walked / in to the lounge / I was grabbed / / that's that's the first time I  
saw my wife / the last

A man had got hold of my wife / / and was threatening her with a knife / / and I was  
/ grabbed from both sides / at the back / and forced down / / he was wearing a /  
clown's mask / / and a / dark blue / boiler type suit

I lost consciousness and didn't / next thing I know I was / on the floor / bound and  
gagged

.... and would urge anybody / that knows anything / about the death of my wife / to  
come forward /

---

#### Case 16

If there's anybody out there who knows anything / just / tell the police or t  
(*unfinished word*) / ever who's done it got a conscience / just give it up you killed  
them

#### Case 17

Well it's hard to sleep really it's just / house don't feel the same without / with /  
them not being there really / you just / just feels empty

Whoever's got Shannon pl (*unfinished word*) just please let her go / her family's  
missing her all her friends are missing her at school. Well I think s (*unfinished word*)  
that somebody out there who knows Shannon / pos (*unfinished word*) pos  
(*unfinished word*) t (*unfinished word*) probably know me as well / and it's just I just  
want her home safe really. It's it makes me think now I can't trust people who's  
really close to me any more / I just can't trust them

---

#### Case 18

She was beautiful

If anyone has seen her can they please contact the police

*Interviewer: can I ask you how how hard this past couple of days has been for a mother who just doesn't know the answer to that question*

No / (not clearly audible)

#### Case 19

...at about eight / recognised that she hadn't awakened me when she came home from running / um / but didn't think much of it she often does that if I'm / tired or something she'll just out of respect / I guess just let me sleep / um / and I often drive her to work today she wanted to drive herself

...then I called her at about ten o'clock / to / just say hi see how she was doing and they told me she / / / sorry

*Unseen person: she never got to work*

She she never made it in this morning

---

#### Case 20

Please come back / I want you back / the the kids wa (*unfinished word*) want you back / kids need you back I need you back / please get in touch / / and please come back

...and if anyone has / seen her / please let us know / where she is /

....there's two small children / / who are g (*unfinished word*) going frantic who are desperate/ for their / mother / thank you

### Case 21

Well what me and Susan would like to say is / we want Louise back / Susan and I are stunned / / Susan is still sedated at the moment / / for christ's sake look after her / whoever's got her / and let us have her back

### Case 22

She's absolutely beautiful / erm / real bubbly / real outgoing real forward / / really friendly / / I love her

Monday morning erm / because I couldn't get / Valentine's day off / / erm / we / we swapped cards / upstairs / erm / / I gave her a / kiss and a cuddle / got ready for work / and she was gonna get her head down for another hour or so / / climbed back in bed / / cuddled up to her again / / / kissed her good bye and went to work /

I love her to bits / I just want her back / she's (cries) / I wanna know where she is /

*Interviewer: What does Joanne mean to you*

---

/ she's the only person I truly love / w (*unfinished* word) wanna know where she is /

/ I want her back

---

### Case 23

Unfortunately it has reached / a point where suspicion of me / is keeping people / from searching for Laci / they've lost that um they're focusing on me / we need to ask people / when was the last time they really thought about Laci missing as

opposed to when / they've thought about the suspicions that swirl round me  
currently / um it's important / that we get people / out there looking for Laci again /  
there's very simple things / um obviously / yes I had a romantic relationship that  
was inappropriate and unfair to a lot of people / and I I apologise / to everyone  
involved in that to all the families / it had nothing to do with Laci's disappearance  
  
Yeah she knows how much she's loved how much she's missed / and be strong and  
/ and we're working to get you home / those are the things /

*Interviewer: you haven't given up*

No god definitely not it's the one thing we have

*Interviewer: how difficult does that get as your baby's due date, this is your first  
(interrupted)*

Every day it's more and more difficult / but / f (*unfinished word*) imagine how  
difficult it must be for her she's / so please / volunteers / anything possible to help  
us / thanks I I c (*unfinished word*) / that'll have to be it Judy / thank you

*Interviewer: Tell me what you love about Laci*

God what's y' know / think you can sum it up / pretty easily with / looking at her  
photograph / the photograph her big smile in the press /

*Interviewer: I have to ask you*

Sure

*Interviewer: Are you in any way connected to Laci's disappearance*

I had nothing to do / with Laci's disappearance

---

#### Case 24

I would like to s (*unfinished word*) to say to whoever / has my children / / that they  
please / I mean please bring 'em home

I wanna say / that your mama loves you so much / and your daddy this whole  
families love you so much / / and you guys have gotta be strong / Hey I mean  
they've got ev (*unfinished word*) people everywhere looking for 'em I mean you  
know you can hear the helicopters up there an / um / but they they haven't seen  
anything / and it's just / it's crazy I I just / and this man since nine o'clock last night  
and you know / and and I really thought that when it got daylight they would find  
'em just in a heartbeat but / all I'm doing is just praying and / and just / keeping my  
faith that they're gonna be okay

---

#### Case 25

.... both Lee and / the other person / were / / like playing cat and mouse with each  
other / for a while / erm / then they / overtook us / I was / shouting at Lee to / you  
know / slow down just ignore them and stop the car but he d (*unfinished word*) / / I  
dunno I dunno if a lot a men are like it and a lot a women are like it but erm / when  
you get behind the wheel of a car / you know sometimes you change / personality /  
it was just the way he looked / he was / his eyes was / he had starey eyes / erm /  
just didn't seem normal

I saw the man / / hit Lee / / I don't know what with / I didn't see anything / um / I  
got out the car / cause I'm not the sort a' person to sit there and er / I got out the  
car / and then I went / over to the er man / / we had a confrontation and / he hit  
me / I can't remember / I fell to the floor I can't remember if I was / erm / / knocked  
out for a bit or what I don't know / I've done a first aid course when I was little /  
erm I was trying to think of everything that I could do / erm / put my coat over him  
and / I didn't move him // and er / I just / / tried to stop / the bleeding really / and /  
comfort him as much as I could / er / driver / er / walked off / it was nothing to do  
with the driver / and all I wanna say is please / will the driver of the car / just come  
forward / because you are not to blame / for this / and I know that / Erm whoever  
this person is that was with you obviously know him / but / / he's ruined my life  
/ he's ruined / the lives / of Maureen and Ray / and Michelle Lee's sister / and Lee's  
little girl Danielle / and erm / please / just tell us who he is because you won't get in  
any trouble at all it was not your fault

---

### **Honest appeals**

#### Case 26

These / recent days have / proven to be the most harrowing of our lives / with / no  
sleep / and / we feel like we're / barely existing / sort of surviving every moment /  
not knowing where Aisling is / is she near us / or has she been moved far away / is  
she being treated well / things like has has her nappy / been changed / just / / these  
thoughts turn through us / as we huddle close / as a family / and we / try to wait /

find out / if there's anything // at the same time the weight of our grief / is  
supported / and shared by all New Zealanders / who have shown an extraordinary  
magnitude of care and empathy / / to the people of New Zealand / who have  
offered their love / support / ke (*unfinished word*) we'd like to say to you / a big  
thank you / particularly to church members / the police / and our community of  
West Auckland /// Aisling is our two year old daughter / a sister to Caitlin / s  
(*unfinished word*) a / a cousin / and a beloved niece / and granddaughter /// we  
want her back / so / please if you have / any information / any / at all / a (*unfinished  
word*) a (*unfinished word*) please come forward / we thank you / thank you all very  
much

#### Case 27

*Interviewer: Okay Carrie / it's been one week / you you have so many people/  
worried / how are you holding up*

I'm / holding up the best that I can / / which isn't very good

*Interviewer: what what is the FBI telling you*

They're they're not / telling me anything / I mean / the police are just just the  
Escondido police department are still head of the / of the um / search

*Interviewer: what is your what is your heart telling you as far as what has happened  
I hate to ask this question Carrie*

I think she's out there somewhere / I just think she's in a bad situation she's trying  
to get out of / and I think she will



### Case 28

Our daughter is the kinda girl that / you dream of / she's the kinda kid that /  
everybody wants as a parent / she's / respectful she's bright she's caring / um she's  
everything that you look for in a great person / um and / she's been that way since  
she showed up / hnh (small smile/laugh sound) / and / I just wanna get her back /  
and everybody here in san diego and / y (*unfinished word*) we're all trying to get her  
back and / we just need more help / let's find her

### Case 29

*Voice off camera (whisper): Take your time*

We appeal / from all our family / anyone who has any information / regarding the  
murder / of my husband Graham / and the attack / on my son / Michael / to come  
forward / / Graham was a wonderful husband / and father / and son / to Pat and  
Ronnie / and brother / to Kenny / Tilly / Terry / Angela and Ian / / and to all our  
family / his loss has left / a huge hole / in all of our lives / Graham was lovely / in  
every way / and he was too young to die / / he had so much to live for / and we had  
so many plans / that have been wiped out / / Graham thought of everybody / and  
how he could help / and he would / / Graham was in the army for twenty three  
years / and loved / loved army life / and in two thousand and five / he left the army  
/ and it was our time / to be a family / he promised he would never leave us / again  
/ / while he was in the army / he also worked in Birkenhead careers / and and in the  
recruitment team / he made a h (*unfinished word*) a lot of impact on / people's lives

that wanted to join the army / he told them / what a great life it was / and when we  
were often in town / people would stop him / and shake his hand and say thanks  
Mac / you you made a difference to our lives / / Graham also served in Northern  
Ireland / and Bosnia / and many other countries / / still cannot believe he's gone /  
how can your loved ones / walk out the front door one Saturday morning / going to  
the match / never to come home again / how can this person / have done this to us  
/ / Graham was my life Graham was my soulmate / h (*unfinished word*) never ever  
stop loving him

*Voice off camera (whisper): Well done*

Thank you

### Case 30

Just want to say / as a parent / you expect to be able to protect / your child from  
everything but / we weren't there / to protect Jo from this / we're we're very /  
we're very angry that someone has taken Jo / from us / and from his family / and  
from / this very large family and we're sure / that any parent / can imagine the pain  
/ we're feeling

### Case 31

There's nothing / that would draw her away from her kids / They ask / why she's  
not here and they ask why she left and then they ask / you know if she didn't leave  
then / / did somebody take her and I / don't have any answers

Keri is / one hundred percent / you know / devoted to her kids you know there is just / nothing / nothing that would have kept her from being there / They've been telling me that / you have to stay positive and it is hard it's extremely hard and / and um / my son's there and my daughter's there and / / (*unintelligible word*) do you have to stay positive / My son is ten and my daughter's eight and / you know they've been asking some pretty tough questions and / I really don't have any answers to give 'em / She is / a soccer mom she's a baseball mom she's a karate mom she / sh (*unfinished word*) they do it two nights in our / every night a week and sometimes two / two different things a night

#### Case 32

To know her was to love her but missing her's gonna be unbearable / I wrote a little note on her mirror in January and it just said best little sister ever / love your big brother / when we cleaned out her apartment yesterday it was still there / I was her big brother you know / / was my job to protect her / and / I was unable to do that

*Voiceover: Kyleigh's older brother, Bernie Sousa, says he won't rest until he finds his sister's killer*

I will never stop / looking for you / I will dedicate every resource I have financially / I do not care what it takes you will be caught / That's all it takes for a case like this is one lead one tip / one thing to change the whole landscape of this case

#### Case 33

My brother was a great guy / and / ehh / h he loved / all people all American people  
/ um / we just / ask anyone who can give us any / inf information / / anything at all /  
/ we'd really appreciate it / um / the American people have been / so welcoming  
and friendly to us / we would like to thank / thank 'em all of them / / for loving  
Lavern / while he was here / / thank you

#### Case 34

You just can't / believe it / you can't / fathom how someone / could have such /  
hate and anger / in them to do / something like this / / oh (distress) / can't /  
understand how s (*unfinished* word) / anybody could do that / / if their family  
member / that had been taken / at such / a manner / they would want / everybody  
possible to / assist the police so please please / just say / it / please / just / say what  
you know if you / see the car

*Voice over of reporter over 'just say it': Marcus Powell, London Tonight*

#### Case 35

We er / have been here / today / to help find Nancy / we believe that she will be  
brought home soon / and we thank the community for the support they've given  
and we / and ask that they continue in their effort to help find our daughter

#### Case 36

Our mum was always so full of life / she was loved by everyone who knew her / /  
we truly miss her / she really was the centre of our family / we can't understand

why anyone would want to hurt her / / our lives will never be the same without her  
/ she was f frail and vulnerable / and the thought of her frightened / / defenceless  
and alone / in her last moments / is the most difficult part / this has distressed all of  
our family / our children have suffered / the loss of a gentle and loving grandma / in  
such a terribly shocking and violent way / / there was no need for my mum's death  
/ she was a generous person / and that if asked for any money she would give 'em  
twice the amount if she could / / and I just would say that if anyone has any  
information to help the investigation / please please come forward

#### Case 37

Rosemary / / please come home / you you haven't done anything wrong / we all I  
(*unfinished word/stammer*) love and miss you terribly / / w (*unfinished  
word/stammer*) / we are a family of five / and the four of us here / need you back  
with us / / our hearts are breaking / please come home / we love you so much

#### Case 38

She's very outgoing/ she's always smiling / she ah plays little jokes on me here and  
there / she um / she likes to do cartwheels / she loves to be outside / she likes to /  
play with her friends / / and I / / I just want her back home

#### Case 39

You can't imagine what Sarah means to us / we're a strong family / and we don't  
survive well apart / we need her home / now / today / as quickly as we possibly can  
/ her brothers and her sister are / they're really not coping very well without her / /

she's our life / / She's a soft / gentle / little girl / she hasn't got a horrible bone in her body / / somebody out there must have seen her / they must have seen her on that road they must have seen her

#### Case 40

We all miss her and love her very much / / contact the police department / anybody who has any information / and whoever has children out there knows how much / / their child means to them / allow them to bring her home safely to us / Shaniya if you're listening to daddy / I miss you so much honey and I'm waiting for you / I'm not gonna give up and you don't give up either honey

#### Case 41

I just want to say how very worried we are about Sian / / /

*Unseen voice: do you want me to do it?*

She's been missing now for over two days / and it's not like her not to come home / or cont (*unfinished word/stammer*) or contact any of us for such a long time / we all want to know where Sian is / a (*unfinished word/stammer*) / and we want her home safe and well / / Sian is a bubbly / lively person / who is instantly liked by everyone / she meets / she's very close to her family and has lots of friends / we'd like to thank all of the people who have shown us support and helped search for Sian / / their kind words and gestures have been a great help to us / / someone out there must have seen / or know where she is / and we just want them to come forward and contact the police / / this is a terrible time for all of us / and we are praying for

Sian's safe return / / if Sian is listening and doesn't want to contact us I beg her to at least ring the police / we're all really grateful for the media coverage which we're sure is helping to spread the appeal / but after the conference today / we would really like to be left alone to deal with this difficult situation / we all just want to know she's okay as it's breaking our hearts not to know where she is

## *Appendix 9*

### **Study 5: Definitions of verbal cues and instructions to coders**

#### **Cues to deception:**

##### Speech errors

Grammatical errors; single word consecutive repetition (eg. 'I I' 'and and'); incomplete/broken sentences; unfinished word/stammer; consecutive phrase repetition (eg. 'when I came when I came to'); mumbling.

Do NOT include: missing 1<sup>st</sup> person pronoun; sentences that don't make sense; single word repetition used for emphasis (eg. 'lovely lovely' 'please please'); normal contractions (eg. 'wanna' 'gonna' ''em')

##### Filled pauses

Non lexical sounds (eg. 'um' 'er' 'ah' 'oh' 'huh')

##### Phrase repetition

Exact phrase repetition (2 or more words). Does not need to be consecutive, so phrase can be anywhere in appeal. Count each phrase as 1 (eg. 'want her back' and 'want her back' counts as 2)

##### Sentence/statement doesn't make sense

Sentence or statement that is illogical, unclear, or doesn't make sense

Do NOT include: incomplete sentences counted in speech errors

##### Statement displaying lack of hope

Expressions which indicate a lack of hope of finding the missing person (eg. 'nothing can be done to make things better', using the past tense about the missing person). Count each phrase/clause so that emphasis is included.

##### Lists



At least 3 items (eg. people, actions, emotions). If list is longer than 3 items, count each 3 items as 1 list.

Do NOT include: positive descriptions of the victim

Avoid 1<sup>st</sup> person pronoun

1<sup>st</sup> person pronoun (singular or plural, ie. 'I' or 'we') disappearing altogether; switching from 1<sup>st</sup> to 2<sup>nd</sup>/3<sup>rd</sup> person pronoun; using general terms instead of 'I' or 'we' (eg. 'everyone's missing her')

Refer to others

References to other people or groups of people.

Do not include: references to police; references to perpetrator; references to family; references to witnesses/as part of plea for help (eg. 'if there's anybody out there who has seen anything' or 'someone must know'); phrases such as 'you know'

Irrelevant statements

Statements of extraneous information, outside context of incident.

Do NOT include: statements about victim

Equivocation

Words or phrases that are: modifiers; minimisers; qualifiers; hedging words; words indicative of uncertainty/not definitive; vague or evasive words (eg. 'just' 'kind of' 'sort of' 'like' 'guess' 'maybe' 'a little' 'probably' 'possibly' 'quite' 'a bit' 'some' 'don't know')

Brutal language or detail

Use of brutal language or detail directly about victim (eg. 'murdered' 'blood')

**Cues to honesty**

Expressions of hope

Expressions which indicate hope of finding the missing person. Count each phrase/clause so that emphasis is included.

Describe victim positively/expression of love

Any positive descriptions of the victim; expressions of love for the victim in 1<sup>st</sup> person ('I' or 'we'); any terms of endearment. Count emphasis words/phrases separately.

Do NOT include: expressions in 2<sup>nd</sup>/3<sup>rd</sup> person

Expression of concern for victim/expressions of pain

Expressions of concern for victim; expression of missing the victim; expressions of wanting victim back (include 1<sup>st</sup> person only eg. 'we want you back' 'I miss her'). Count emphasis words/phrases separately (eg. 'we need her back, now, today' counts as 3; 'we all miss and love her very much' counts as 3); expressions of grief or pain

Do NOT include: circumlocutions (eg. 'she knows how much she's missed' 'all parents know how much their child means to them'); expressions in 2<sup>nd</sup> or 3<sup>rd</sup> person (eg. 'her family's missing her' or 'you just miss them' or 'everyone's grieving')

Plea

Any request for assistance: plea for help from public/witnesses (eg. 'somebody must have seen her'); direct plea to victim (eg. 'please come home'). Count each phrase/clause so that emphasis is included (eg. 'please, please come back' counts as 3).

Avoid brutal language

Avoiding use of brutal language in relation to victim (eg. 'gone' or 'taken from us' or 'done this' instead of 'murdered')

Reference to norms of emotion/behaviour (including violations of norms)

Questions/lack of understanding/disbelief about how somebody could do this/why somebody would do this/how this could happen; reference to normal emotional standards (eg. 'any parent knows')

## ***Appendix 10***

### **Study 6: Definitions of nonverbal cues and instructions to coders**

#### **Instructions for coding data**

##### **Gaze aversion**

Appealer not looking directly at the camera/s or interviewer, but looking down, to the side or up. Wearing sunglasses or a hat so that eyes cannot be seen is not included as a form of gaze aversion.

Please note down (using the time display on the playback program) each individual instance of gaze aversion, including when the gaze aversion begins and when it ends.

Please time to the nearest half second.

##### **Shrugging**

Please count all instances of shrugging (including shrugging both shoulders, or just one shoulder). Please note down, using the time display on the playback program, when each shrug occurs.

##### **Head shaking**

Please count all instances of head shaking (some may be subtle), counting each individual head shake when there is a cluster (ie several head shakes together).

Please note down, using the time display on the playback program, when each head shake occurs.

## *Appendix 11*

### **Study 7: Participant response sheet**

#### **Participant instructions**

Thank you for agreeing to participate in this study, which is investigating behaviours ('cues') associated with deception in a high stakes situation.

You will be shown a number of short video clips of a type shown in general news programmes. The clips show people talking about missing or murdered relatives, some of whom are honest and some of whom are deceptive. In the clips in which you can see more than one person, please focus on the speaker. All the clips were filmed shortly after the relatives went missing or were murdered (between one day and one month later).

Before viewing any of the clips you will be asked to estimate what percentage of the appeals you think you will correctly judge as truthful or deceptive.

On the response sheet, for each clip there is a brief summary of what was publicly known of the case at the time of the appeal. The summary will explain any events, details or names mentioned by the person in the clip. Please read the summary before watching the clip. After watching each clip, you will be asked to note on the response sheet whether you think the person featured in the clip is deceptive or honest, and to rate how confident you are in your decision.

On the response sheet is also a box to check if you are familiar with the person featured in the clip, or the outcome of the case. Please do not complete the response sheet for the clip if you have checked this box.

If there is anything that you do not understand in the clip, for example if any words spoken are unclear, please ask the researcher. This process will be repeated for each clip.

After viewing all the clips, you will be asked to estimate what percentage of the appeals that you think you correctly judged as truthful or deceptive. You will then be asked to list any cues that you have used to come to your decisions. You will then be given a list of cues that may be associated with honesty or deception, and asked to note how often you use them when making your decisions.

**What percentage of clips do you think you will correctly judge as truthful or deceptive?**

---

**Clip 1**

**This clip features a woman talking about her missing daughter, Morgan, who is 20 years old.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident* *Reasonably confident* *Very confident*  
1..... 2..... 3..... 4..... 5.....

---

**Clip 2**

**This clip features a man talking about his missing wife, Leah.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident* *Reasonably confident* *Very confident*  
1..... 2..... 3..... 4..... 5.....

---

**Clip 3**

**This clip features a woman talking about her missing daughter, Karissa, who is 12 years old.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

Not at all confident                      Reasonably confident                      Very confident  
 1..... 2..... 3..... 4..... 5.....

#### **Clip 4**

**This clip features a man talking about his murdered daughter, Tania, who was 19 years old.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      Lying                      Truthful

How confident are you in your decision (please circle a number):

Not at all confident                      Reasonably confident                      Very confident  
 1..... 2..... 3..... 4..... 5.....

#### **Clip 5**

**This clip features a man talking about his murdered wife. She was found stabbed in the front garden of their home.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      Lying                      Truthful

How confident are you in your decision (please circle a number):

Not at all confident                      Reasonably confident                      Very confident  
 1..... 2..... 3..... 4..... 5.....

#### **Clip 6**

**This clip features a man talking about his murdered brother, Lavern.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      Lying                      Truthful

How confident are you in your decision (please circle a number):

Not at all confident                      Reasonably confident                      Very confident  
 1..... 2..... 3..... 4..... 5.....

#### **Clip 7**

**This clip features a woman talking about her murdered husband, Graham. It was reported that Graham and their teenage son, Michael, had been attacked.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

---

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

#### **Clip 8**

**This clip features a woman talking about her murdered children, Devon and Damon. Darin is her husband. She reported that she had been attacked and her children murdered by an intruder in the family home. The clip was filmed on what would have been Devon's 7<sup>th</sup> birthday.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

#### **Clip 9**

**This clip features a man talking about his missing 34 year old daughter, Nancy.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

#### **Clip 10**

**This clip features a man talking about his missing fiancé, Joanne.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

#### **Clip 11**

**This clip features a man talking about his missing girlfriend, Sian.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

---

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

**Clip 12**

**This clip features a woman talking about her missing 9 year old daughter, Shannon.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 13**

**This clip features a man talking about his missing wife.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 14**

**This clip features a man talking about his missing 15 year old daughter, Rosemary.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 15**

**This clip features a man talking about the murder of his fiancé. He reported that he and his fiancé had been robbed in their car and his fiancé had been shot.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*



---

How confident are you in your decision (please circle a number):

<i>Not at all confident</i>			<i>Reasonably confident</i>			<i>Very confident</i>
1.....	2.....	3.....	4.....	5.....		

#### **Clip 16**

**This clip features a man talking about his murdered sister, Kyleigh, who was 21 years old. She was found robbed and murdered.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

<i>Not at all confident</i>			<i>Reasonably confident</i>			<i>Very confident</i>
1.....	2.....	3.....	4.....	5.....		

---

#### **Clip 17**

**This clip features a woman talking about her missing 6 year old daughter.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

<i>Not at all confident</i>			<i>Reasonably confident</i>			<i>Very confident</i>
1.....	2.....	3.....	4.....	5.....		

---

#### **Clip 18**

**This clip features a man talking about his missing 2 year old stepson. He reported that he himself had been beaten up, and his step-son had been kidnapped.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

<i>Not at all confident</i>			<i>Reasonably confident</i>			<i>Very confident</i>
1.....	2.....	3.....	4.....	5.....		

---

#### **Clip 19**

**This clip features a woman talking about her missing 8 year old daughter, Sarah.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident* *Reasonably confident* *Very confident*  
1..... 2..... 3..... 4..... 5.....

---

**Clip 20**

**This clip features a man talking about his pregnant, missing wife, Laci.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident* *Reasonably confident* *Very confident*  
1..... 2..... 3..... 4..... 5.....

---

**Clip 21**

**This clip features a man talking about his missing 17 year old daughter.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident* *Reasonably confident* *Very confident*  
1..... 2..... 3..... 4..... 5.....

---

**Clip 22**

**This clip features a woman talking about her missing 8 year old daughter.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is: *Lying* *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident* *Reasonably confident* *Very confident*  
1..... 2..... 3..... 4..... 5.....

---

**Clip 23**

**This clip features a man talking about his murdered wife. He reported that intruders had bound and gagged him, and taken his wife to gain entry to the bank where she worked.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

---

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 24**

**This clip features a man talking about his 23 year old missing daughter, Linda.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 25**

**This clip features a woman talking about her murdered 18 year old son.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 26**

**This clip features a woman talking about her two missing sons, who were 1 and 3 years old. She reported that she had been carjacked by a man, who had driven away with her two sons in the back seat.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 27**

**This clip features a man talking about his missing 5 year old daughter, Shaniya.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

---

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 28**

**This clip features a man talking about his missing wife.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 29**

**This clip features a woman talking about the murder of her fiancé, Lee. She reported that he had been murdered after a road rage incident.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 30**

**This clip features a woman talking about her murdered mother.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 31**

**This clip features a woman, Carrie, talking about her missing 14 year old daughter.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

---

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 32**

**This clip features a man talking about his missing, newborn daughter, Louise. Susan is his wife.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 33**

**This clip features a man talking about his missing daughter, Aisling, who is 2 years old.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 34**

**This clip features a man talking about his missing wife, Keri.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*

1..... 2..... 3..... 4..... 5.....

---

**Clip 35**

**This clip features a man talking about his dead children. His 3 young children died in an arson attack on their home.**

Please check this box and do not complete the response sheet for this clip, if you are ☐

---

familiar with the person featured in the clip or the outcome of the case:

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*  
1..... 2..... 3..... 4..... 5.....

**Clip 36**

**This clip features a man talking about his murdered son, Joe, who was 20 years old.**

Please check this box and do not complete the response sheet for this clip, if you are familiar with the person featured in the clip or the outcome of the case: ☐

Do you think the person in the clip is:                      *Lying*                      *Truthful*

How confident are you in your decision (please circle a number):

*Not at all confident*                      *Reasonably confident*                      *Very confident*  
1..... 2..... 3..... 4..... 5.....  
...                      ...                      ...                      ...                      ...

---

**What percentage of clips do you think you judged correctly?**

**What verbal or nonverbal cues did you use most in this study to decide whether the people in the clips were lying or telling the truth?**

**Of the cues you have listed, which do you consider the most important?**

Did you use any of the following cues in this study to decide whether the people in the clips were lying or telling the truth?

For each cue, please indicate how often you used it in this study.

If you used the cue, please indicate whether you thought it was related to honesty or deception.

**Cues to do with the voice and speech of the appealers**

The appealers were hesitant when talking:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

The appealers were certain and direct in what they were saying:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

The appealers were speaking in a personal way, and sounded expressive and involved in what they were saying:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

The appealers' voice was quivering with genuine emotion:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

The appealers' voice was high-pitched:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

The appealers used words or phrases which made them sound tentative, vague, uncertain or evasive (for example, 'sort of' 'kind of' 'a bit' 'quite' 'maybe' 'just' 'like' 'don't know' 'I guess' 'possibly'):

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*



---

The appellant produced speech errors (for example, unfinished words or sentences, consecutive word or phrase repetition, mumbling, stammering, grammatical errors, sentences that don't make sense):

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

The appellant avoided using brutal language about their relative (for example, saying 'taken from us' or 'gone' instead of 'murdered'):

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

The appellant expressed concern for their relative, missing their relative, or wanting their relative back:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

The appellant described their relative positively, talked about their relative in positive terms, expressed love for their relative, or used terms of endearment about their relative:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

The appellant referred to normal emotional behaviour (eg. 'any parent knows how much their child means to them') or to violations of normal behaviour (eg. 'how could anybody do this'):

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

### **Cues to do with the emotions of the appellants**

The appellant's emotions were fake:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

The appellant's facial expressions were fake:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	

---

The appealer seemed genuinely sad:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appealer's eyes looked genuinely sad:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appealer seemed emotionally cold:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appeal seemed genuine and heartfelt:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appealer seemed urgent:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
In cases where the relative is missing, the appealer seemed hopeful of finding them:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appealer seemed defensive:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appealer was showing no emotional variation:				
<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>	<i>Deception</i>	
The appealer was trying to stay calm:				

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer was trying to contain their emotions:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

**Cues to do with the behaviour of the appealers**

The appealer seemed to be putting on a performance:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer was pretending to cry:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer seemed plausible:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer was behaving unnaturally:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer seemed like a normal person:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer was looking away a lot/did not have direct eye contact:

<i>Never</i>	<i>Rarely</i>	<i>Sometimes</i>	<i>Often</i>	<i>Very often</i>
When I used the cue, I thought it indicated:		<i>Honesty</i>		<i>Deception</i>

---

The appealer had credible eye contact (for example, if they were looking down a lot, it was because they were reading a statement):

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

---

**Cues to do with your personal reaction to the appealers**

The appealer seemed creepy:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

---

I could feel the appealer's pain:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

---

I did NOT feel sympathy for the appealer:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

---

I disliked the appealer:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

---

I felt sorry for the appealer:

*Never*

*Rarely*

*Sometimes*

*Often*

*Very often*

When I used the cue, I thought it indicated:

*Honesty*

*Deception*

---